Addendum to:
Model Specification Testing of Time Series Regressions

1. Introduction

In this addendum to Bierens (1984) [B84 hereafter], I will derive the limiting null distribution of the proposed Weighted Integrated Condition Moment (WICM) test involved and show how to derive upper bounds of the critical values of a standardized version of the WICM test. Moreover, I will show how to approximate the critical values of the WICM test via a bootstrap method. Furthermore, I will give a more formal proof of the consistency of the WICM test. Also, I will solve the problem how to standardize the conditioning variables before applying a bounded transformation such that all the asymptotic results carry over. This was one of the unsolved problems in B84.

The finite sample power of the WICM will be demonstrated via a numerical example, and in the last section I will sketch avenues for optimizing the finite sample power.

As in B84, I will focus on the nonlinear ARX model

$$
\begin{aligned}
Y_t &= f(Y_{t-1}, Y_{t-2}, ..., Y_{t-p}, X_{t-1}, X_{t-2}, ..., X_{t-q}, \theta_0) + U_t && (1.1)\\
&= f_{t-1}(\theta_0) + U_t, \text{ say,} && (1.2)
\end{aligned}
$$

where (1.2) is merely a short-hand notation for (1.1), and

Assumption 1.1.
(a) $Z_t = (Y_t, X_t')' \in \mathbb{R} \times \mathbb{R}^s$ *is a strictly stationary vector time series process, with* $E[Y_t^2] < \infty$ *and* X_t *is a vector of exogenous variables*;
(b) θ_0 *is a parameter vector contained in a given compact parameter space* $\Theta \subset \mathbb{R}^m$;
(c) $f(v, \theta)$ *is a given real function on* $\mathbb{R}^{p+q.s} \times \Theta$ *which for each* $v \in \mathbb{R}^{p+q.s}$ *is continuous in* $\theta \in \Theta$, *and for each* $\theta \in \Theta$ *Borel measurable in* $v \in \mathbb{R}^{p+q.s}$, *hence for each* $\theta \in \Theta$ *the sequence* $f_{t-1}(\theta)$ *is a well-defined strictly stationary time series process*;
(d) $\sup_{\theta\in\Theta} E[f_{t-1}(\theta)^2] < \infty$,

and U_t is the error term.

In general, time series regression models aim to represent the best one-step ahead forecasting scheme, in the sense that given the entire past of the time series involved up to time $t-1$, the mean square forecast error is minimal. For the model (1.1) this is the case if and only if the error process U_t is a martingale difference sequence w.r.t. the σ-algebra

$$\mathcal{F}_{-\infty}^{t} = \sigma\left(\{Z_{t-j}\}_{j=0}^{\infty}\right)$$

generated by the sequence $\{Z_{t-j}\}_{j=0}^{\infty}$, i.e.,

$$U_t \text{ is measurable } \mathcal{F}_{-\infty}^{t} \text{ and } E\left[U_t|\mathcal{F}_{-\infty}^{t-1}\right] = 0 \text{ a.s.}$$

More formally, using the short-hand notation (1.2), the null hypothesis to be tested is that

$$H_0: \text{ There exists a } \theta_0 \in \Theta \text{ such that } \Pr\left(E\left[Y_t|\mathcal{F}_{-\infty}^{t-1}\right] = f_{t-1}(\theta_0)\right) = 1,$$

which is equivalent to the hypothesis that $U_t = Y_t - f_{t-1}(\theta_0)$ is a martingale difference process, against the alternative hypothesis that H_0 is false, i.e.,

$$H_1: \text{ For all } \theta \in \Theta,\ \Pr\left(E\left[Y_t|\mathcal{F}_{-\infty}^{t-1}\right] = f_{t-1}(\theta)\right) < 1.$$

Note that if we define θ_0 as

$$\begin{aligned}
\theta_0 &= \arg\min_{\theta\in\Theta} E\left[(Y_t - f_{t-1}(\theta))^2\right] \qquad (1.3)\\
&= \arg\min_{\theta\in\Theta} \left\{E\left[\left(Y_t - E\left[Y_t|\mathcal{F}_{-\infty}^{t-1}\right]\right)^2\right] + E\left[\left(E\left[Y_t|\mathcal{F}_{-\infty}^{t-1}\right] - f_{t-1}(\theta)\right)^2\right]\right\}\\
&= \arg\min_{\theta\in\Theta} E\left[\left(E\left[Y_t|\mathcal{F}_{-\infty}^{t-1}\right] - f_{t-1}(\theta)\right)^2\right]
\end{aligned}$$

regardless whether H_0 is true or not, then H_0 and H_1 become

$$H_0 : \Pr\left(E\left[Y_t|\mathcal{F}_{-\infty}^{t-1}\right] = f_{t-1}(\theta_0)\right) = 1, \qquad (1.4)$$
$$H_1 : \Pr\left(E\left[Y_t|\mathcal{F}_{-\infty}^{t-1}\right] = f_{t-1}(\theta_0)\right) < 1, \qquad (1.5)$$

respectively.

To determine which one of the two hypotheses is true, we need to condition on the one-sided infinite vector time series $\{Z_{t-j}\}_{j=1}^{\infty}$, which in practice is not possible because Z_t is usually only observed from a particular time t_0 onwards. Part of the solution to this problem is the following lemma.

Lemma 1.1. *Under Assumption 1.1,*

$$E\left[Y_t|\mathcal{F}_{-\infty}^{t-1}\right] = \lim_{k\to\infty} E\left[Y_t|\mathcal{F}_{t-k}^{t-1}\right] \ a.s., \tag{1.6}$$

where $\mathcal{F}_{t-k}^{t-1} = \sigma\left(\{Z_{t-j}\}_{j=1}^{k}\right)$ *is the* σ*-algebra generated by* $Z_{t-1}, Z_{t-2}, ..., Z_{t-k}$. *Consequently, under* H_1 *and with* θ_0 *defined by* (1.3) *there exists an* $k_0 \in \mathbb{N}$ *such that for all* t,

$$\sup_{k\geq k_0} \Pr\left(E\left[Y_t|\mathcal{F}_{t-k}^{t-1}\right] = f_{t-1}(\theta_0)\right) < 1. \tag{1.7}$$

Proof. The result (1.6) is well-known. See for example Theorem 9.4.8 in Chung (1974) or Theorem 3.12 in Bierens (2004). As to part (1.7), let

$$V_t = \left|E\left[Y_t|\mathcal{F}_{-\infty}^{t-1}\right] - f_{t-1}(\theta_0)\right|, \ V_{t,k} = \left|E\left[Y_t|\mathcal{F}_{t-k}^{t-1}\right] - f_{t-1}(\theta_0)\right|.$$

Then by (1.6), $\lim_{k\to\infty} V_{t,k} = V_t$ a.s., which implies that $V_{t,k} \xrightarrow{\mathrm{d}} V_t$ for $k \to \infty$ and thus, by definition of convergence in distribution,

$$\lim_{k\to\infty} \Pr\left[V_{t,k} > \varepsilon\right] = \Pr\left[V_t > \varepsilon\right] \tag{1.8}$$

for all continuity points ε of the distribution function of V_t. Moreover, under the alternative hypothesis (1.5), $\Pr\left[V_t > 0\right] > 0$, which implies that there exists a continuity point $\varepsilon > 0$ of the distribution function of V_t such that $\Pr\left[V_t > \varepsilon\right] > 0$ as well.[1] It follows now from (1.8) that

$$\liminf_{k\to\infty} \Pr\left[V_{t,k} > 0\right] \geq \lim_{k\to\infty} \Pr\left[V_{t,k} > \varepsilon\right] = \Pr\left[V_t > \varepsilon\right] > 0,$$

hence

$$\limsup_{k\to\infty} \Pr\left(E\left[Y_t|\mathcal{F}_{t-k}^{t-1}\right] = f_{t-1}(\theta_0)\right) = \limsup_{k\to\infty} \Pr\left[V_{t,k} = 0\right] < 1. \tag{1.9}$$

Since by the strict stationarity condition in Assumption 1.1 the probabilities in (1.9) do not depend on t, the result (1.7) follows. ■

Now let $\Phi : \mathbb{R}^{s+1} \to \mathbb{R}^{s+1}$ be a bounded Borel measurable one-to-one mapping with Borel measurable inverse. Then

$$\mathcal{F}_{t-k}^{t-1} = \sigma\left(\{Z_{t-j}\}_{j=1}^{k}\right) = \sigma\left(\{\Phi(Z_{t-j})\}_{j=1}^{k}\right),$$

[1] Because $\Pr\left[V_t > 1/n\right] \uparrow \Pr\left[V_t > 0\right]$ monotonically as $n \to \infty$.

hence by Theorem 1 in Bierens (1982) and its generalization in Theorem 2.1 in the addendum to Bierens (1982) in Chapter 2,

$$H_1(k) : \Pr\left(E\left[Y_t|\mathcal{F}_{t-k}^{t-1}\right] = f_{t-1}(\theta_0)\right) < 1$$

implies that the set

$$\begin{aligned} S_k &= \Bigg\{\tau = (\tau_1', \tau_2', ...\tau_k')' \in \mathbb{R}^{(s+1)k} : \\ &\qquad E\left[(Y_t - f_{t-1}(\theta_0))\exp\left(\mathbf{i}.\sum_{j=1}^{k}\tau_j'\Phi(Z_{t-j})\right)\right] = 0\Bigg\} \end{aligned}$$

has Lebesgue measure zero and is nowhere dense. Consequently, denoting

$$\psi_k(\tau_1, \tau_2, ...\tau_k) = E\left[(Y_t - f_{t-1}(\theta_0))\exp\left(\mathbf{i}.\sum_{j=1}^{k}\tau_j'\Phi(Z_{t-j})\right)\right] \tag{1.10}$$

the following result holds.

Lemma 1.2. *Assume that the conditions of Lemma 1.1 hold. Let Υ be a compact subset of $\mathbb{R}^{s+1}$ with positive Lebesgue measure and let μ be an absolutely continuous probability measure on Υ. Then*

$$\int_{\Upsilon^k} |\psi_k(\tau_1, \tau_2, ...\tau_k)|^2 \,\mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k) \begin{cases} > 0 \text{ under } H_1(k), \\ = 0 \text{ under } H_0. \end{cases}$$

Given that Z_t is observed for $1 - t_0 \leq t \leq n$, where $t_0 = \max(p, q)$, the empirical counter-part of $\psi_k(\tau_1, \tau_2, ...\tau_k)$ defined in (1.10) is

$$\widetilde{\psi}_{k,n}(\tau_1, \tau_2, ...\tau_k) = \frac{1}{n}\sum_{t=1}^{n}(Y_t - f_{t-1}(\widehat{\theta}_n))\exp\left(\mathbf{i}.\sum_{j=1}^{\min(k,t-t_0)}\tau_j'\Phi(Z_{t-j})\right) \tag{1.11}$$

for $n > k + t_0$, where

$$\widehat{\theta}_n = \arg\min_{\theta\in\Theta}\frac{1}{n}\sum_{t=1}^{n}(Y_t - f_{t-1}(\theta))^2$$

is the NLLS estimator of θ_0.

However, for the time being I will assume that all the lagged Z_t's are observed, so that

$$\widehat{\psi}_{k,n}(\tau_1, \tau_2, ...\tau_k) = \frac{1}{n}\sum_{t=1}^{n}(Y_t - f_{t-1}(\widehat{\theta}_n)) \exp\left(\mathbf{i}.\sum_{j=1}^{k}\tau_j'\Phi(Z_{t-j})\right) \tag{1.12}$$

will be treated as the empirical counter-part of $\psi_k(\tau_1, \tau_2, ...\tau_k)$. In section 7 I will set forth conditions under which the asymptotic results below on the basis of (1.12) are the same as for (1.11).

In order to keep this addendum in tune with the addendums to Bierens (1982), Bierens and Ploberger (1997) and Bierens and Wang (2012) in Chapters 2, 5 and 6, respectively, I will from this point onwards denote $\sqrt{n}\widehat{\psi}_{k,n}(\tau_1, \tau_2, ...\tau_k)$ by $\widehat{W}_{k,n}(\tau_1, \tau_2, ...\tau_k)$. Thus,

$$\widehat{W}_{k,n}(\tau_1, \tau_2, ...\tau_k) = \frac{1}{\sqrt{n}}\sum_{t=1}^{n}(Y_t - f_{t-1}(\widehat{\theta}_n)) \exp\left(\mathbf{i}.\sum_{j=1}^{k}\tau_j'\Phi(Z_{t-j})\right). \tag{1.13}$$

In this addendum to B84 I will set forth further conditions such that under H_0 the following results hold.

- For each $k \in \mathbb{N}$,

$$\begin{aligned}\widehat{B}_{n,k} &\stackrel{\text{def.}}{=} \int_{\Upsilon^k}\left|\widehat{W}_{k,n}(\tau_1, \tau_2, ...\tau_k)\right|^2 \mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k) \qquad (1.14)\\ &\stackrel{\mathrm{d}}{\to} B_k \stackrel{\text{def.}}{=} \int_{\Upsilon^k}\left|W_k(\tau_1, \tau_2, ...\tau_k)\right|^2 \mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k),\end{aligned}$$

  where $W_k(\tau_1, \tau_2, ...\tau_k)$ is a complex-valued zero-mean Gaussian process on Υ^k;

- for any positive sequence γ_k satisfying $\sum_{k=1}^{\infty}\gamma_k < \infty$ and any subsequence L_n of n satisfying $\lim_{n\to\infty} L_n = \infty$,

$$\sum_{k=1}^{L_n}\gamma_k\widehat{B}_{n,k} \stackrel{\mathrm{d}}{\to} \sum_{k=1}^{\infty}\gamma_k B_k;$$

- the critical values of $\sum_{k=1}^{\infty}\gamma_k B_k$ can be approximated by a bootstrap method, similar to the addendum to Bierens (1982) in Chapter 2.

Moreover, under H_1 the following results hold.

- For each $k \in \mathbb{N}$,

$$\widehat{B}_{n,k}/n \xrightarrow{\mathrm{p}} \eta_k \overset{\text{def.}}{=} \int_{\Upsilon^k} |\psi_k(\tau_1, \tau_2, ...\tau_k)|^2 \,\mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k) \qquad (1.15)$$

  where $\eta_k > 0$ for all but a finite number of k's,

- and

$$\frac{1}{n}\sum_{k=1}^{L_n} \gamma_k \widehat{B}_{n,k} \xrightarrow{\mathrm{p}} \sum_{k=1}^{\infty} \gamma_k \eta_k > 0$$

The statistic

$$\widehat{T}_n = \sum_{k=1}^{L_n} \gamma_k \widehat{B}_{n,k}$$

will be referred to as the test statistic of the WICM test.

2. Nonlinear least squares

2.1. Consistency

In B84 the consistency of the NLLS estimator $\widehat{\theta}_n$ is derived under data heterogeneity. However, under the current strict stationarity condition in Assumption 1.1 the following conditions suffice.

Assumption 2.1.
(a) *The vector time series process Z_t in Assumption 1.1(a) has a vanishing memory, in the sense that the sets in its remote σ-algebra $\mathcal{F}_{-\infty} = \cap_t \mathcal{F}_{-\infty}^t$ have either probability zero or one;*
(b) $E[\sup_{\theta\in\Theta} f_{t-1}(\theta)^2] < \infty$;
(c) *The solution θ_0 of* (1.3) *is unique.*

See Bierens (2004, Ch. 7) for the motivation of the vanishing memory concept. In particular this concept plays a key-role in the following uniform weak law of large numbers (UWLLN).

Lemma 2.1. *Let $V_t \in \mathbb{R}^k$ be a strictly stationary time series process with vanishing memory, and let Θ be a compact subset of a Euclidean space. Let $g(v,\theta)$ be a real or complex-valued function on $\mathbb{R}^k \times \Theta$ satisfying the following conditions.*
(1) *For each $\theta \in \Theta$, $g(v,\theta)$ is Borel measurable in $v \in \mathbb{R}^k$;*
(2) *For each $v \in \mathbb{R}^k$, $g(v,\theta)$ is continuous in $\theta \in \Theta$;*
(3) $E\left[\sup_{\theta\in\Theta}|g(V_1,\theta)|\right] < \infty$.
Then $p\lim_{n\to\infty}\sup_{\theta\in\Theta}\left|\frac{1}{n}\sum_{t=1}^n g(V_t,\theta) - E[g(V_1,\theta)]\right| = 0$.

Proof. Similar to Theorem 7.8(a) in Bierens (2004, p. 187). ■

This lemma mimics the uniform strong law of large number of Jennrich (1969), with Kolmogorov's strong law of large numbers replaced by the following weak version for time series.

Lemma 2.2. *Let $V_t \in \mathbb{R}^k$ be a strictly stationary time series process with vanishing memory, with $E[||V_t||] < \infty$. Then*

$$p\lim_{n\to\infty}\frac{1}{n}\sum_{t=1}^n V_t = E[V_1].$$

Proof. See Bierens (2004, Theorem 7.4, p. 184). ■

Denoting $V_t = (Z_t', Z_{t-1}', ..., Z_{t-\max(p,q)}')'$, it follows from Assumptions 1.1 and 2.1(a) that V_t is strictly stationary with vanishing memory.[2] Next, let

$$g(V_t,\theta) = (Y_t - f_{t-1}(\theta))^2$$

and note that by Assumptions 1.1 and 2.1(b),

$$\begin{aligned} E\left[\sup_{\theta\in\Theta}|g(V_1,\theta)|\right] &= E\left[\sup_{\theta\in\Theta}(Y_t - f_{t-1}(\theta))^2\right] \\ &\le 2.E[Y_t^2] + 2.E\left[\sup_{\theta\in\Theta} f_{t-1}(\theta)^2\right] < \infty. \end{aligned}$$

Therefore, it follows from Lemma 2.1 that under Assumption 1.1 and parts (a) and (b) of Assumption 2.1,

$$p\lim_{n\to\infty}\sup_{\theta\in\Theta}\left|\widehat{Q}_n(\theta) - Q(\theta)\right| = 0, \tag{2.1}$$

[2]Because $\mathcal{F}_{-\infty}^t = \sigma\left(\{Z_{t-j}\}_{j=0}^{\infty}\right) = \sigma\left(\{V_{t-j}\}_{j=0}^{\infty}\right)$.

where

$$\widehat{Q}_n(\theta) = \frac{1}{n}\sum_{t=1}^{n}(Y_t - f_{t-1}(\theta))^2\,,\ Q(\theta) = E\left[(Y_t - f_{t-1}(\theta))^2\right]. \tag{2.2}$$

It is now a standard nonlinear regression exercise[3] to verify that

Lemma 2.1. *Under Assumptions 1.1 and 2.1 the NLLS estimator $\widehat{\theta}_n$ satisfies $p\lim_{n\to\infty}\widehat{\theta}_n = \theta_0$, regardless whether H_0 is true or not.*

Of course, if H_0 is not true then $U_t = Y_t - f_{t-1}(\theta_0)$ is no longer a martingale difference process.

2.2. Asymptotic normality

In addition to Assumptions 1.1 and 2.1 the following conditions are sufficient for the asymptotic normality of $\sqrt{n}(\widehat{\theta}_n - \theta_0)$.

Assumption 2.2. *Under H_0 the following conditions hold.*
(a) Θ is convex and θ_0 is an interior point of Θ;
(b) $f_{t-1}(\theta)$ is a.s. twice continuously differentiable in the components $\theta_1, \theta_2, ..., \theta_m$ of θ;
(c) for $i_1, i_2 = 1, 2, ..., m$,

$$\begin{aligned} E\left[(Y_t - f_{t-1}(\theta))^2.\left|(\partial/\partial\theta_{i_1})f_{t-1}(\theta)\right|.\left|(\partial/\partial\theta_{i_2})f_{t-1}(\theta)\right|\Big|_{\theta=\theta_0}\right] &< \infty, & (2.3)\\ E\left[\sup_{\theta\in\Theta}\left|(\partial/\partial\theta_{i_1})f_{t-1}(\theta)\right|.\left|(\partial/\partial\theta_{i_2})f_{t-1}(\theta)\right|\right] &< \infty, & (2.4)\\ E\left[\sup_{\theta\in\Theta}\left|Y_t - f_{t-1}(\theta)\right|.\left|(\partial/\partial\theta_{i_1})(\partial/\partial\theta_{i_2})f_{t-1}(\theta)\right|\right] &< \infty; & (2.5) \end{aligned}$$

(d) The matrix $A_2 = E\left[((\partial/\partial\theta')f_{t-1}(\theta))\,((\partial/\partial\theta)f_{t-1}(\theta))|_{\theta=\theta_0}\right]$ is nonsingular.

Denoting

$$\nabla f_{t-1}(\theta) = (\partial/\partial\theta')f_{t-1}(\theta)$$

[3] See for example Bierens (2004, Ch. 6).

it follows then from the first-order conditions for a minimum of $\widehat{Q}_n(\theta)$ and the mean value theorem that, with probability converging to 1,

$$0 = \frac{1}{\sqrt{n}} \sum_{j=1}^{n} U_t . \nabla f_{t-1}(\theta_0) - \widetilde{A}_{2,n} \sqrt{n} (\widehat{\theta}_n - \theta_0)$$

where

$$\begin{aligned} \widetilde{A}_{2,n} &= \frac{1}{n} \sum_{j=1}^{n} \begin{pmatrix} ((\partial/\partial\theta_1) f_{t-1}(\theta)) \, (\nabla f_{t-1}(\theta))'|_{\theta=\widetilde{\theta}_1} \\ \vdots \\ ((\partial/\partial\theta_m) f_{t-1}(\theta)) \, (\nabla f_{t-1}(\theta))'|_{\theta=\widetilde{\theta}_m} \end{pmatrix} \\ &\quad - \frac{1}{n} \sum_{j=1}^{n} \begin{pmatrix} (Y_t - f_{t-1}(\theta)) \, ((\partial/\partial\theta_1) (\nabla f_{t-1}(\theta))')|_{\theta=\widetilde{\theta}_1} \\ \vdots \\ (Y_t - f_{t-1}(\theta)) \, ((\partial/\partial\theta_m) (\nabla f_{t-1}(\theta))')|_{\theta=\widetilde{\theta}_m} \end{pmatrix}, \end{aligned}$$

with the $\widetilde{\theta}_i$'s mean values satisfying

$$||\widetilde{\theta}_i - \theta_0|| \leq ||\widehat{\theta}_n - \theta_0||, \; i = 1, 2, ..., m. \tag{2.6}$$

Since by Assumption 2.2(c) and Lemma 2.1,

$$\begin{aligned} p \lim_{n\to\infty} \sup_{\theta\in\Theta} \bigg| \frac{1}{n} \sum_{j=1}^{n} \big(((\partial/\partial\theta_{i_1}) f_{t-1}(\theta)) \, ((\partial/\partial\theta_{i_2}) f_{t-1}(\theta)) \\ - E\left[((\partial/\partial\theta_{i_1}) f_{t-1}(\theta)) \, ((\partial/\partial\theta_{i_2}) f_{t-1}(\theta)) \right] \Big) \bigg| = 0 \end{aligned}$$

and

$$\begin{aligned} p \lim_{n\to\infty} \sup_{\theta\in\Theta} \bigg| \frac{1}{n} \sum_{j=1}^{n} \big((Y_t - f_{t-1}(\theta)) \, ((\partial/\partial\theta_{i_i}) (\partial/\partial\theta_{i_2}) f_{t-1}(\theta)) \\ - E\left[(Y_t - f_{t-1}(\theta)) \, ((\partial/\partial\theta_i) (\partial/\partial\theta_{i_2}) f_{t-1}(\theta)) \right] \Big) \bigg| = 0, \end{aligned}$$

it follows straightforwardly from (2.6) and Lemma 2.1 that $p \lim_{n\to\infty} \widetilde{A}_{2,n} = A_2$, hence by Assumption 2.2(d), $p \lim_{n\to\infty} \widetilde{A}_{2,n}^{-1} = A_2^{-1}$.

Finally, note that $U_t.\nabla f_{t-1}(\theta_0)$ is a vector-valued martingale difference process, so that by the martingale difference central limit theorem of McLeish (1974) and condition (2.3),

$$\frac{1}{\sqrt{n}}\sum_{t=1}^{n} U_t.\nabla f_{t-1}(\theta_0) \xrightarrow{\mathrm{d}} N_m[0, A_1], \text{ where}$$
$$A_1 = E\left[U_t^2.(\nabla f_{t-1}(\theta_0))(\nabla f_{t-1}(\theta_0))'\right].$$

Summarizing, the following asymptotic normality result has been shown to hold.

Lemma 2.2. *Under H_0 and Assumptions 1.1, 2.1 and 2.2,*

$$\begin{aligned}\sqrt{n}(\widehat{\theta}_n - \theta_0) &= \frac{1}{\sqrt{n}}\sum_{j=1}^{n} U_t.A_2^{-1}\nabla f_{t-1}(\theta_0) + o_p(1) \\ &\xrightarrow{\mathrm{d}} N_m\left[0, A_2^{-1}A_1A_2^{-1}\right].\end{aligned}$$

3. The limiting null distribution of the WICM test

I will first show that

Lemma 3.1. *Under H_0 and Assumptions 1.1, 2.1 and 2.2, $\widehat{B}_{n,k} \leq 2.\widehat{B}_{1,n,k} + 2\widehat{B}_{2,n,k}$, where $E[\widehat{B}_{1,n,k}] = E[U_1^2]$ and $\sup_{k\in\mathbb{N}} \widehat{B}_{2,n,k} = O_p(1)$, hence for any positive sequence γ_k satisfying $\sum_{k=1}^{\infty}\gamma_k < \infty$ and any subsequence L_n of n such that $\lim_{n\to\infty} L_n = \infty$,*

$$\sum_{k=1}^{\infty}\gamma_k\widehat{B}_{n,k} - \sum_{k=1}^{L_n}\gamma_k\widehat{B}_{n,k} = o_p(1). \tag{3.1}$$

Proof. It follows from (1.13) that

$$\begin{aligned}&\widehat{W}_{k,n}(\tau_1, \tau_2, ...\tau_k) \\ &\quad = \frac{1}{\sqrt{n}}\sum_{t=1}^{n} U_t \exp\left(\mathbf{i}.\sum_{j=1}^{k}\tau_j'\Phi(Z_{t-j})\right) \\ &\qquad - \frac{1}{\sqrt{n}}\sum_{t=1}^{n}\left(f_{t-1}(\widehat{\theta}_n) - f_{t-1}(\theta_0)\right)\exp\left(\mathbf{i}.\sum_{j=1}^{k}\tau_j'\Phi(Z_{t-j})\right)\end{aligned} \tag{3.2}$$

so that

$$\begin{aligned}
\widehat{B}_{n,k} &= \int_{\Upsilon^k} \left|\widehat{W}_{k,n}(\tau_1,\tau_2,....\tau_k)\right|^2 \mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k) \\
&= \int_{\Upsilon^k} \left(\frac{1}{\sqrt{n}}\sum_{t=1}^{n} U_t \cos\left(\sum_{j=1}^{k} \tau_j'\Phi(Z_{t-j})\right)\right. \\
&\quad \left. -\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\left(f_{t-1}(\widehat{\theta}_n) - f_{t-1}(\theta_0)\right)\cos\left(\sum_{j=1}^{k} \tau_j'\Phi(Z_{t-j})\right)\right)^2 \\
&\quad \times \mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k) \\
&\quad + \int_{\Upsilon^k} \left(\frac{1}{\sqrt{n}}\sum_{t=1}^{n} U_t \sin\left(\sum_{j=1}^{k} \tau_j'\Phi(Z_{t-j})\right)\right. \\
&\quad \left. -\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\left(f_{t-1}(\widehat{\theta}_n) - f_{t-1}(\theta_0)\right)\sin\left(\sum_{j=1}^{k} \tau_j'\Phi(Z_{t-j})\right)\right)^2 \\
&\quad \times \mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k) \\
&\leq 2\int_{\Upsilon^k} \left(\frac{1}{\sqrt{n}}\sum_{t=1}^{n} U_t \cos\left(\sum_{j=1}^{k} \tau_j'\Phi(Z_{t-j})\right)\right)^2 \mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k) \\
&\quad +2\int_{\Upsilon^k} \left(\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\left(f_{t-1}(\widehat{\theta}_n) - f_{t-1}(\theta_0)\right)\cos\left(\sum_{j=1}^{k} \tau_j'\Phi(Z_{t-j})\right)\right)^2 \\
&\quad \times \mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k) \\
&\quad +2\int_{\Upsilon^k} \left(\frac{1}{\sqrt{n}}\sum_{t=1}^{n} U_t \sin\left(\sum_{j=1}^{k} \tau_j'\Phi(Z_{t-j})\right)\right)^2 \mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k) \\
&\quad +2\int_{\Upsilon^k} \left(\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\left(f_{t-1}(\widehat{\theta}_n) - f_{t-1}(\theta_0)\right)\sin\left(\sum_{j=1}^{k} \tau_j'\Phi(Z_{t-j})\right)\right)^2 \\
&= 2.\widehat{B}_{1,n,k} + 2\widehat{B}_{2,n,k}
\end{aligned}$$

where

$$\widehat{B}_{1,n,k} = \int_{\Upsilon^k} \left|\frac{1}{\sqrt{n}}\sum_{t=1}^{n} U_t \exp\left(\mathbf{i}.\sum_{j=1}^{k} \tau_j'\Phi(Z_{t-j})\right)\right|^2 \mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k)$$

$$
\begin{aligned}
&= \int_{\Upsilon^k} \left(\frac{1}{\sqrt{n}} \sum_{t=1}^{n} U_t \cos\left(\sum_{j=1}^{k} \tau_j' \Phi(Z_{t-j}) \right) \right)^2 \mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k) \\
&\quad + \int_{\Upsilon^k} \left(\frac{1}{\sqrt{n}} \sum_{t=1}^{n} U_t \sin\left(\sum_{j=1}^{k} \tau_j' \Phi(Z_{t-j}) \right) \right)^2 \mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k), \\
\widehat{B}_{2,n,k} &= \int_{\Upsilon^k} \left| \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \left(f_{t-1}(\widehat{\theta}_n) - f_{t-1}(\theta_0) \right) \exp\left(\mathbf{i}. \sum_{j=1}^{k} \tau_j' \Phi(Z_{t-j}) \right) \right|^2 \\
&\quad \times \mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k) \\
&= \int_{\Upsilon^k} \left(\frac{1}{\sqrt{n}} \sum_{t=1}^{n} \left(f_{t-1}(\widehat{\theta}_n) - f_{t-1}(\theta_0) \right) \cos\left(\sum_{j=1}^{k} \tau_j' \Phi(Z_{t-j}) \right) \right)^2 \\
&\quad \times \mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k) \\
&\quad + \int_{\Upsilon^k} \left(\frac{1}{\sqrt{n}} \sum_{t=1}^{n} \left(f_{t-1}(\widehat{\theta}_n) - f_{t-1}(\theta_0) \right) \sin\left(\sum_{j=1}^{k} \tau_j' \Phi(Z_{t-j}) \right) \right) \\
&\quad \times \mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k).
\end{aligned}
$$

Obviously,

$$E[\widehat{B}_{1,n,k}] = E[U_t^2] = E[U_1^2].$$

To prove that $\sup_{k\in\mathbb{N}} \widehat{B}_{2,n,k} = O_p(1)$, observe that by the mean value theorem,

$$
\begin{aligned}
&\frac{1}{\sqrt{n}} \sum_{t=1}^{n} \left(f_{t-1}(\widehat{\theta}_n) - f_{t-1}(\theta_0) \right) \cos\left(\sum_{j=1}^{k} \tau_j' \Phi(Z_{t-j}) \right) \\
&= \sqrt{n}(\widehat{\theta}_n - \theta_0)' \frac{1}{n} \sum_{t=1}^{n} \nabla f_{t-1}(\widetilde{\theta}_n(\tau)) \cos\left(\sum_{j=1}^{k} \tau_j' \Phi(Z_{t-j}) \right)
\end{aligned} \tag{3.3}
$$

where $\widetilde{\theta}_n(\tau)$ is a mean value such that $||\widetilde{\theta}_n(\tau) - \theta_0|| \leq ||\widehat{\theta}_n - \theta_0||$, hence

$$
\begin{aligned}
&\left| \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \left(f_{t-1}(\widehat{\theta}_n) - f_{t-1}(\theta_0) \right) \cos\left(\sum_{j=1}^{k} \tau_j' \Phi(Z_{t-j}) \right) \right| \\
&\qquad \leq ||\sqrt{n}(\widehat{\theta}_n - \theta_0)||. \frac{1}{n} \sum_{t=1}^{n} \sup_{\theta\in\Theta} ||\nabla f_{t-1}(\theta)||
\end{aligned}
$$

and thus,

$$\int_{\Upsilon^k}\left(\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\left(f_{t-1}(\widehat{\theta}_n)-f_{t-1}(\theta_0)\right)\cos\left(\sum_{j=1}^{k}\tau_j'\Phi(Z_{t-j})\right)\right)^2$$
$$\times\, \mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k)$$
$$\leq ||\sqrt{n}(\widehat{\theta}_n-\theta_0)||.\frac{1}{n}\sum_{t=1}^{n}\sup_{\theta\in\Theta}||\nabla f_{t-1}(\theta)||$$
$$= O_p(1),\ \text{uniformly in } k.$$

The latter follows from the fact that by Lemma 2.2, $||\sqrt{n}(\widehat{\theta}_n-\theta_0)|| = O_p(1)$ and by condition (2.4), $E[\sup_{\theta\in\Theta}||\nabla f_{t-1}(\theta)||^2] < \infty$, hence $\frac{1}{n}\sum_{t=1}^{n}\sup_{\theta\in\Theta}||\nabla f_{t-1}(\theta)|| = O_p(1)$.

Similarly,

$$\int_{\Upsilon^k}\left(\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\left(f_{t-1}(\widehat{\theta}_n)-f_{t-1}(\theta_0)\right)\sin\left(\sum_{j=1}^{k}\tau_j'\Phi(Z_{t-j})\right)\right)^2$$
$$\times\, \mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k)$$
$$\leq ||\sqrt{n}(\widehat{\theta}_n-\theta_0)||.\frac{1}{n}\sum_{t=1}^{n}\sup_{\theta\in\Theta}||\nabla f_{t-1}(\theta)||$$
$$= O_p(1),\ \text{uniformly in } k.$$

Hence

$$\sup_{k\in\mathbb{N}}\widehat{B}_{2,n,k} \leq 2||\sqrt{n}(\widehat{\theta}_n-\theta_0)||.\frac{1}{n}\sum_{t=1}^{n}\sup_{\theta\in\Theta}||\nabla f_{t-1}(\theta)|| = O_p(1).$$

The result (3.1) follows now trivially from the fact that $\sum_{k=1}^{\infty}\gamma_k < \infty$ implies $\lim_{\ell\to\infty}\sum_{k=\ell+1}^{\infty}\gamma_k = 0$. ■

Next, denote

$$b_k(\tau_1,\tau_2,....\tau_k) = E\left[\nabla f_{t-1}(\theta_0)\exp\left(\mathbf{i}\sum_{j=1}^{k}\tau_j'\Phi(Z_{t-j})\right)\right], \tag{3.4}$$
$$\phi_{k,t-1}(\tau_1,\tau_2,....\tau_k) = \exp\left(\mathbf{i}\sum_{j=1}^{k}\tau_j'\Phi(Z_{t-j})\right)$$

$$
\begin{aligned}
& & -b_k(\tau_1,\tau_2,....\tau_k)' A_2^{-1} \nabla f_{t-1}(\theta_0), & \quad (3.5)\\
W_{k,n}(\tau_1,\tau_2,....\tau_k) &=& \frac{1}{\sqrt{n}} \sum_{t=1}^{n} U_t . \phi_{k,t-1}(\tau_1,\tau_2,....\tau_k), & \quad (3.6)\\
B_{k,n} &=& \int_{\Upsilon^k} |W_{k,n}(\tau_1,\tau_2,....\tau_k)|^2 \, \mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k). &
\end{aligned}
$$

Then

Lemma 3.2. *Under H_0 and Assumptions 1.1, 2.1 and 2.2, and for fixed $k \in \mathbb{N}$,*

$$
\sup_{(\tau_1,\tau_2,....\tau_k)\in\Upsilon^k} \left| \widehat{W}_{k,n}(\tau_1,\tau_2,....\tau_k) - W_{k,n}(\tau_1,\tau_2,....\tau_k) \right| = o_p(1), \tag{3.7}
$$

hence,

$$
\widehat{B}_{n,k} = B_{k,n} + o_p(1). \tag{3.8}
$$

Proof. Observe from (3.3) that

$$
\begin{aligned}
&\left| \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \left(f_{t-1}(\widehat{\theta}_n) - f_{t-1}(\theta_0) \right) \cos\left(\sum_{j=1}^{k} \tau_j' \Phi(Z_{t-j}) \right) \right.\\
&\qquad \left. - \sqrt{n}(\widehat{\theta}_n - \theta_0)' \frac{1}{n} \sum_{t=1}^{n} \nabla f_{t-1}(\theta_0) \cos\left(\sum_{j=1}^{k} \tau_j' \Phi(Z_{t-j}) \right) \right|\\
&\quad \leq ||\sqrt{n}(\widehat{\theta}_n - \theta_0)|| \frac{1}{n} \sum_{t=1}^{n} \sup_{||\theta-\theta_0||\leq||\widehat{\theta}_n-\theta_0||} ||\nabla f_{t-1}(\theta) - \nabla f_{t-1}(\theta_0)||\\
&\quad = o_p(1) \qquad (3.9)
\end{aligned}
$$

because $||\sqrt{n}(\widehat{\theta}_n - \theta_0)|| = O_p(1)$ and

$$
\frac{1}{n} \sum_{t=1}^{n} \sup_{||\theta-\theta_0||\leq||\widehat{\theta}_n-\theta_0||} ||\nabla f_{t-1}(\theta) - \nabla f_{t-1}(\theta_0)|| = o_p(1). \tag{3.10}
$$

To prove the latter, observe that for an arbitrary small $\varepsilon > 0$,

$$
E\left[\frac{1}{n} \sum_{t=1}^{n} \sup_{||\theta-\theta_0||\leq||\widehat{\theta}_n-\theta_0||} ||\nabla f_{t-1}(\theta) - \nabla f_{t-1}(\theta_0)|| \right]
$$

$$
\begin{aligned}
&\leq E\left[\sup_{||\theta-\theta_0||\leq\varepsilon}||\nabla f_{t-1}(\theta)-\nabla f_{t-1}(\theta_0)||\times I\left(||\widehat{\theta}_n-\theta_0||\leq\varepsilon\right)\right]\\
&\quad+2E\left[\sup_{\theta\in\Theta}||\nabla f_{t-1}(\theta)||\times I\left(||\widehat{\theta}_n-\theta_0||>\varepsilon\right)\right]\\
&\leq E\left[\sup_{||\theta-\theta_0||\leq\varepsilon}||\nabla f_{t-1}(\theta)-\nabla f_{t-1}(\theta_0)||\right]\\
&\quad+2\sqrt{E\left[\sup_{\theta\in\Theta}||\nabla f_{t-1}(\theta)||^2\right]}\sqrt{\Pr\left[||\widehat{\theta}_n-\theta_0||>\varepsilon\right]}\\
&=E\left[\sup_{||\theta-\theta_0||\leq\varepsilon}||\nabla f_{t-1}(\theta)-\nabla f_{t-1}(\theta_0)||\right]+o(1), \qquad (3.11)
\end{aligned}
$$

where the $o(1)$ term is due to the facts that by Lemma 2.1, $\Pr[||\widehat{\theta}_n-\theta_0||>\varepsilon]\to 0$ and that by condition (2.4), $E[\sup_{\theta\in\Theta}||\nabla f_{t-1}(\theta)||^2]<\infty$. Moreover, since

$$
\begin{aligned}
E\left[\sup_{||\theta-\theta_0||\leq\varepsilon}||\nabla f_{t-1}(\theta)-\nabla f_{t-1}(\theta_0)||\right] &\leq 2E\left[\sup_{\theta\in\Theta}||\nabla f_{t-1}(\theta)||\right]\\
&\leq \sqrt{E\left[\sup_{\theta\in\Theta}||\nabla f_{t-1}(\theta)||^2\right]}<\infty
\end{aligned}
$$

and by a.s. continuity of $\nabla f_{t-1}(\theta)$,

$$
\lim_{\varepsilon\downarrow 0}\sup_{||\theta-\theta_0||\leq\varepsilon}||\nabla f_{t-1}(\theta)-\nabla f_{t-1}(\theta_0)||=0 \text{ a.s.},
$$

it follows from the dominated convergence theorem that

$$
\lim_{\varepsilon\downarrow 0}E\left[\sup_{||\theta-\theta_0||\leq\varepsilon}||\nabla f_{t-1}(\theta)-\nabla f_{t-1}(\theta_0)||\right]=0. \qquad (3.12)
$$

It follows now straightforwardly from (3.11) and (3.12) that (3.10) holds.

Next, observe from Lemma 2.1 and Assumption 2.1(b) that for fixed k,

$$
\sup_{(\tau_1,\tau_2,....\tau_k)\in\Upsilon^k}\left\|\frac{1}{n}\sum_{t=1}^{n}\nabla f_{t-1}(\theta_0)\cos\left(\sum_{j=1}^{k}\tau_j'\Phi(Z_{t-j})\right)-\mathrm{Re}[b_k(\tau_1,\tau_2,....\tau_k)]\right\|=o_p(1)
$$

and thus,

$$\sqrt{n}(\widehat{\theta}_n-\theta_0)'\frac{1}{n}\sum_{t=1}^{n}\nabla f_{t-1}(\theta_0)\cos\left(\sum_{j=1}^{k}\tau_j'\Phi(Z_{t-j})\right)$$
$$=\sqrt{n}(\widehat{\theta}_n-\theta_0)'\operatorname{Re}[b_k(\tau_1,\tau_2,....\tau_k)]+o_p(1)$$
$$=\operatorname{Re}[b_k(\tau_1,\tau_2,....\tau_k)]'A_2^{-1}\frac{1}{\sqrt{n}}\sum_{j=1}^{n}U_t.\nabla f_{t-1}(\theta_0)+o_p(1),$$

uniformly on Υ^k, where the latter equality follows from Lemma 2.2. Hence by (3.2) and (3.9),

$$\operatorname{Re}\left[\widehat{W}_{k,n}(\tau_1,\tau_2,....\tau_k)\right]=\operatorname{Re}\left[\frac{1}{\sqrt{n}}\sum_{t=1}^{n}U_t\exp\left(\mathbf{i}.\sum_{j=1}^{k}\tau_j'\Phi(Z_{t-j})\right)\right]$$
$$-\operatorname{Re}[b_k(\tau_1,\tau_2,....\tau_k)]'A_2^{-1}\frac{1}{\sqrt{n}}\sum_{j=1}^{n}U_t.\nabla f_{t-1}(\theta_0))+o_p(1)$$
$$=\operatorname{Re}\left[\frac{1}{\sqrt{n}}\sum_{t=1}^{n}U_t\phi_{k,t-1}(\tau_1,\tau_2,....\tau_k)\right]+o_p(1)$$
$$=\operatorname{Re}\left[W_{k,n}(\tau_1,\tau_2,....\tau_k)\right]+o_p(1)$$

uniformly on Υ^k. The same applies to the Im[.] case. This proves (3.7).

The proof that the latter implies (3.8) is not hard[4] and is therefore left to the reader. ■

I will now set forth conditions such that $\sup_{k\in\mathbb{N}}E[B_{k,n}]<\infty$, as follows. Observe from (3.4), (3.5) and (3.6) that

$$E\left[|W_{k,n}(\tau_1,\tau_2,....\tau_k)|^2\right]$$
$$=E\left[U_t^2\phi_{k,t-1}(\tau_1,\tau_2,....\tau_k)\overline{\phi_{k,t-1}(\tau_1,\tau_2,....\tau_k)}\right]$$

[4] Denote the left-hand side of (3.7) by $d_{k,n}$, and verify that

$$\left|\widetilde{B}_{k,n}-\widehat{B}_{k,n}\right|\le d_{k,n}^2+4.d_{k,n}\sqrt{\widehat{B}_{k,n}}.$$

$$
\begin{aligned}
&= E\left[U_t^2\left(\exp\left(\mathbf{i}\sum_{j=1}^{k}\tau_j'\Phi(Z_{t-j})\right) - b_k(\tau_1,\tau_2,....\tau_k)'A_2^{-1}\nabla f_{t-1}(\theta_0)\right)\right.\\
&\quad\left.\times\left(\exp\left(-\mathbf{i}\sum_{j=1}^{k}\tau_j'\Phi(Z_{t-j})\right) - (\nabla f_{t-1}(\theta_0))'A_2^{-1}\overline{b_k(\tau_1,\tau_2,....\tau_k)}\right)\right]\\
&= E[U_t^2]\\
&- b_k(\tau_1,\tau_2,....\tau_k)'A_2^{-1}E\left[U_t^2\nabla f_{t-1}(\theta_0)\exp\left(-\mathbf{i}\sum_{j=1}^{k}\tau_j'\Phi(Z_{t-j})\right)\right]\\
&- E\left[U_t^2(\nabla f_{t-1}(\theta_0))'\exp\left(\mathbf{i}\sum_{j=1}^{k}\tau_j'\Phi(Z_{t-j})\right)\right]A_2^{-1}\overline{b_k(\tau_1,\tau_2,....\tau_k)}\\
&+ b_k(\tau_1,\tau_2,....\tau_k)'A_2^{-1}E\left[U_t^2\nabla f_{t-1}(\theta_0)\nabla f_{t-1}(\theta_0))'\right]A_2^{-1}\overline{b_k(\tau_1,\tau_2,....\tau_k)}\\
&= E[U_1^2]\\
&\qquad - b_k(\tau_1,\tau_2,....\tau_k)'A_2^{-1}\overline{c_k(\tau_1,\tau_2,....\tau_k)}\\
&\qquad - c_k(\tau_1,\tau_2,....\tau_k)'A_2^{-1}\overline{b_k(\tau_1,\tau_2,....\tau_k)}\\
&\qquad + b_k(\tau_1,\tau_2,....\tau_k)'A_2^{-1}A_1A_2^{-1}\overline{b_k(\tau_1,\tau_2,....\tau_k)},
\end{aligned}
$$

where the bar denotes the complex-conjugate, and

$$
c_k(\tau_1,\tau_2,....\tau_k) = E\left[U_t^2\nabla f_{t-1}(\theta_0)\exp\left(\mathbf{i}\sum_{j=1}^{k}\tau_j'\Phi(Z_{t-j})\right)\right].
$$

It is an easy complex calculus exercise to verify that

$$
\begin{aligned}
&b_k(\tau_1,\tau_2,....\tau_k)'A_2^{-1}\overline{c_k(\tau_1,\tau_2,....\tau_k)}\\
&\qquad + c_k(\tau_1,\tau_2,....\tau_k)'A_2^{-1}\overline{b_k(\tau_1,\tau_2,....\tau_k)}\\
&\quad = 2\operatorname{Re}\left[b_k(\tau_1,\tau_2,....\tau_k)'\right]A_2^{-1}\operatorname{Re}\left[c_k(\tau_1,\tau_2,....\tau_k)\right]\\
&\qquad + 2\operatorname{Im}\left[b_k(\tau_1,\tau_2,....\tau_k)'\right]A_2^{-1}\operatorname{Im}\left[c_k(\tau_1,\tau_2,....\tau_k)\right]
\end{aligned}
$$

and

$$
\begin{aligned}
&b_k(\tau_1,\tau_2,....\tau_k)'A_2^{-1}A_1A_2^{-1}\overline{b_k(\tau_1,\tau_2,....\tau_k)}\\
&\quad = \operatorname{Re}\left[b_k(\tau_1,\tau_2,....\tau_k)'\right]A_2^{-1}A_1A_2^{-1}\operatorname{Re}\left[b_k(\tau_1,\tau_2,....\tau_k)\right]\\
&\qquad + \operatorname{Im}\left[b_k(\tau_1,\tau_2,....\tau_k)'\right]A_2^{-1}A_1A_2^{-1}\operatorname{Im}\left[b_k(\tau_1,\tau_2,....\tau_k)\right],
\end{aligned}
$$

hence

$$
\begin{aligned}
E\left[|W_{k,n}(\tau_1,\tau_2,....\tau_k)|^2\right] &\leq E[U_1^2] \\
&+ 2\,\mathrm{Re}\left[b_k(\tau_1,\tau_2,....\tau_k)'\right] A_2^{-1}\,\mathrm{Re}\left[c_k(\tau_1,\tau_2,....\tau_k)\right] \\
&+ 2\,\mathrm{Im}\left[b_k(\tau_1,\tau_2,....\tau_k)'\right] A_2^{-1}\,\mathrm{Im}\left[c_k(\tau_1,\tau_2,....\tau_k)\right] \\
&+ \mathrm{Re}\left[b_k(\tau_1,\tau_2,....\tau_k)'\right] A_2^{-1}A_1A_2^{-1}\,\mathrm{Re}\left[b_k(\tau_1,\tau_2,....\tau_k)\right] \\
&+ \mathrm{Im}\left[b_k(\tau_1,\tau_2,....\tau_k)'\right] A_2^{-1}A_1A_2^{-1}\,\mathrm{Im}\left[b_k(\tau_1,\tau_2,....\tau_k)\right] \qquad (3.13)
\end{aligned}
$$

Let $\lambda_1,...,\lambda_m$ be the eigenvalues of A_2^{-1}, with corresponding orthonormal eigenvectors $q_1,...,q_m$. Then

$$
\begin{aligned}
&\left|\mathrm{Re}\left[b_k(\tau_1,\tau_2,....\tau_k)'\right] A_2^{-1}\,\mathrm{Re}\left[c_k(\tau_1,\tau_2,....\tau_k)\right]\right| \\
&\leq \sum_{i=1}^{m}\lambda_i\left|q_i'\,\mathrm{Re}\left[q_i'b_k(\tau_1,\tau_2,....\tau_k)\right]\right|.\left|\mathrm{Re}\left[q_i'c_k(\tau_1,\tau_2,....\tau_k)\right]\right| \\
&= \sum_{i=1}^{m}\lambda_i\left|E\left[(\nabla f_{t-1}(\theta_0))'q_i\cos\left(\sum_{j=1}^{k}\tau_j'\Phi(Z_{t-j})\right)\right]\right| \\
&\qquad\times\left|E\left[U_t^2(\nabla f_{t-1}(\theta_0))'q_i\cos\left(\sum_{j=1}^{k}\tau_j'\Phi(Z_{t-j})\right)\right]\right| \\
&\leq \sum_{i=1}^{m}\lambda_i E\left[\left|(\nabla f_{t-1}(\theta_0))'q_i\right|\right]E\left[U_t^2\left|(\nabla f_{t-1}(\theta_0))'q_i\right|\right] \\
&\leq \sum_{i=1}^{m}\lambda_i E\left[||\nabla f_{t-1}(\theta_0)||\right]E\left[U_t^2||\nabla f_{t-1}(\theta_0)||\right] \\
&\leq \sum_{i=1}^{m}\lambda_i\sqrt{E\left[||\nabla f_{t-1}(\theta_0)||^2\right]}E\left[U_t^2||\nabla f_{t-1}(\theta_0)||\right] \\
&= \sum_{i=1}^{m}\lambda_i\sqrt{\mathrm{trace}(A_2)}E\left[U_t^2||\nabla f_{t-1}(\theta_0)||\right] \\
&= \mathrm{trace}(A_2^{-1})\sqrt{\mathrm{trace}(A_2)}E\left[U_t^2||\nabla f_{t-1}(\theta_0)||\right]
\end{aligned}
$$

and similarly

$$
\begin{aligned}
&\mathrm{Im}\left[b_k(\tau_1,\tau_2,....\tau_k)'\right] A_2^{-1}\,\mathrm{Im}\left[c_k(\tau_1,\tau_2,....\tau_k)\right] \\
&\qquad\leq \mathrm{trace}(A_2^{-1})\sqrt{\mathrm{trace}(A_2)}E\left[U_t^2||\nabla f_{t-1}(\theta_0)||\right]
\end{aligned}
$$

Moreover, with $\lambda_1^*, ..., \lambda_m^*$ the eigenvalues of $A_2^{-1} A_1 A_2^{-1}$ we have similarly,

$$\begin{aligned}
&\operatorname{Re}\left[b_k(\tau_1, \tau_2, ...\tau_k)'\right] A_2^{-1} A_1 A_2^{-1} \operatorname{Re}\left[b_k(\tau_1, \tau_2, ...\tau_k)\right] \\
&\quad \leq \sum_{i=1}^{m} \lambda_i^* \left(E\left[||\nabla f_{t-1}(\theta_0)||\right]\right)^2 \\
&\quad \leq \sum_{i=1}^{m} \lambda_i^* E\left[||\nabla f_{t-1}(\theta_0)||^2\right] \\
&\quad = \sum_{i=1}^{m} \lambda_i^* .\text{trace}(A_2) \\
&\quad = \text{trace}(A_2^{-1} A_1 A_2^{-1}).\text{trace}(A_2)
\end{aligned}$$

and

$$\begin{aligned}
&\operatorname{Im}\left[b_k(\tau_1, \tau_2, ...\tau_k)'\right] A_2^{-1} A_1 A_2^{-1} \operatorname{Im}\left[b_k(\tau_1, \tau_2, ...\tau_k)\right] \\
&\quad \leq \text{trace}(A_2^{-1} A_1 A_2^{-1}).\text{trace}(A_2).
\end{aligned}$$

Hence,

$$\begin{aligned}
E\left[B_{k,n}\right] &= \int_{\Upsilon^k} E\left[\left|W_{k,n}(\tau_1, \tau_2, ...\tau_k)\right|^2\right] \mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k) \\
&\leq E[U_1^2] + 4.\text{trace}(A_2^{-1})\sqrt{\text{trace}(A_2)} E\left[U_t^2 ||\nabla f_{t-1}(\theta_0)||\right] \\
&\quad +2.\text{trace}(A_2^{-1} A_1 A_2^{-1}).\text{trace}(A_2). \qquad (3.14)
\end{aligned}$$

Thus, if

Assumption 3.1. $E\left[U_t^2 ||\nabla f_{t-1}(\theta_0)||\right] < \infty$

then

Lemma 3.3. *Under H_0 and Assumptions 1.1, 2.1, 2.2 and 3.1,* $\sup_{k \in \mathbb{N}} E[B_{k,n}] < \infty$.

Lemmas 3.1-3.3 now yield the following corollary.

Lemma 3.4. *Under H_0 and Assumptions 1.1, 2.1, 2.2 and 3.1,*

$$\sum_{k=1}^{\infty} \gamma_k \widehat{B}_{k,n} = \sum_{k=1}^{\infty} \gamma_k B_{k,n} + o_p(1)$$

for any positive sequence γ_k *satisfying* $\sum_{k=1}^{\infty}\gamma_k < \infty$.

Proof. For any $K \in \mathbb{N}$,

$$\left|\sum_{k=1}^{\infty}\gamma_k\widehat{B}_{k,n} - \sum_{k=1}^{\infty}\gamma_k B_{k,n}\right| \leq R_{0,n,K} + R_{1,n,K} + R_{2,n,K}$$

where

$$\begin{aligned} R_{0,n,K} &= \left|\sum_{k=1}^{K}\gamma_k\widehat{B}_{k,n} - \sum_{k=1}^{K}\gamma_k B_{k,n}\right|, \\ R_{1,n,K} &= 2\sum_{k=K+1}^{\infty}\gamma_k\widehat{B}_{1,k,n} + \sum_{k=K+1}^{\infty}\gamma_k B_{k,n}, \\ R_{2,n,K} &= 2\left(\sum_{k=K+1}^{\infty}\gamma_k\right)\sup_{k\in\mathbb{N}}\widehat{B}_{2,k,n}. \end{aligned}$$

Let $\varepsilon > 0$ and $\delta > 0$ be arbitrary. By Chebyshev's inequality for first moments and Lemmas 3.1 and 3.3 there exists a $K_1(\delta,\varepsilon) \in \mathbb{N}$ such that

$$\begin{aligned} \Pr\left[R_{1,n,K} > \delta/2\right] &\leq 2E[R_{1,n,K}]/\delta \\ &= 2\frac{2\sum_{k=K+1}^{\infty}\gamma_k E[\widehat{B}_{1,k,n}] + \sum_{k=K+1}^{\infty}\gamma_k E[B_{k,n}]}{\delta} \\ &\leq \delta^{-1}\left(\sum_{k=K+1}^{\infty}\gamma_k\right)\left(4E[U_1] + 2\sup_{k\in\mathbb{N}}E[B_{k,n}]\right) \\ &< \varepsilon \end{aligned}$$

for all $K \geq K_1(\delta,\varepsilon)$. Moreover, by Lemma 3.1 there exists a $K_2(\delta,\varepsilon) \in \mathbb{N}$ such that

$$\Pr\left[R_{2,n,K} > \delta/2\right] = \Pr\left[\sup_{k\in\mathbb{N}}\widehat{B}_{2,k,n} > \frac{\delta}{4\sum_{k=K+1}^{\infty}\gamma_k}\right] < \varepsilon$$

for all $K \geq K_2(\delta,\varepsilon)$. Using the easy inequality

$$\Pr[R_1 + R_2 > \delta] \leq \Pr[R_1 > \delta/2] + \Pr[R_2 > \delta/2] \tag{3.15}$$

for nonnegative random variables R_1 and R_2, it follows now that for a fixed $K_0 \geq \max\{K_1(\delta,\varepsilon), K_2(\delta,\varepsilon)\}$,

$$\Pr\left[R_{1,n,K_0} + R_{2,n,K_0} > \delta\right] < 2\varepsilon$$

Finally, it follows from Lemma 3.2 that $p\lim_{n\to\infty} R_{0,n,K_0} = 0$, so that there exists an $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$,

$$\Pr\left[R_{2,n,K_0} > \delta\right] < \varepsilon.$$

Hence by (3.15),

$$\begin{aligned}
\Pr&\left[\left|\sum_{k=1}^{\infty} \gamma_k \widehat{B}_{k,n} - \sum_{k=1}^{\infty} \gamma_k B_{k,n}\right| > 2\delta\right] \\
&\leq \Pr\left[R_{2,n,K_0} + R_{1,n,K_0} + R_{2,n,K_0} > 2\delta\right] \\
&\leq \Pr\left[R_{2,n,K_0} + R_{1,n,K_0} + R_{2,n,K_0} > 2\delta\right] \\
&\leq \Pr\left[R_{2,n,K_0} > \delta\right] + \Pr\left[R_{1,n,K_0} + R_{2,n,K_0} > \delta\right] \\
&< 3\varepsilon,
\end{aligned}$$

which implies that $p\lim_{n\to\infty}\left|\sum_{k=1}^{\infty}\gamma_k \widehat{B}_{k,n} - \sum_{k=1}^{\infty}\gamma_k B_{k,n}\right| = 0.$ ■

The next step is to prove that for each $k \in \mathbb{N}$, $W_{k,n}(\tau_1, \tau_2, ...\tau_k)$ converges weakly to a zero-mean complex-valued Gaussian process $W_k(\tau_1, \tau_2, ...\tau_k)$ on Υ^k, denoted by $W_{k,n} \Rightarrow W_k$, by showing that $W_{k,n}(\tau_1, \tau_2, ...\tau_k)$ is tight on Υ^k and its finite distributions converge to the corresponding finite distribution of $W_k(\tau_1, \tau_2, ...\tau_k)$. See the addendum to Bierens (1982) in Chapter 2 and Billingsley (1968) for these concepts. However, it follows from Bierens and Ploberger (1997, Lemma A.1) that $\mathrm{Re}[W_{k,n}(\tau_1, \tau_2, ...\tau_k)]$ and $\mathrm{Im}[W_{k,n}(\tau_1, \tau_2, ...\tau_k)]$ are tight on Υ^k, hence so is $W_{k,n}(\tau_1, \tau_2, ...\tau_k)$. Now it suffices to verify that the finite distributions of $W_{k,n}(\tau_1, \tau_2, ...\tau_k)$ converge to the corresponding finite distribution of $W_k(\tau_1, \tau_2, ...\tau_k)$, which can be done similar to the addendum to Bierens (1982) in Chapter 2, replacing the reference to the standard central limit theorem for the i.i.d. case with a reference to the martingale difference central limit theorem of McLeish (1974). Thus,

Lemma 3.5. *Under H_0 and Assumptions 1.1, 2.1 and 2.2, and for each $k \in \mathbb{N}$, $W_{k,n}(\tau_1, \tau_2, ..., \tau_k) \Rightarrow W_k(\tau_1, \tau_2, ..., \tau_k)$ on Υ^k, where the latter is a zero-mean* *complex-valued Gaussian process on Υ^k with covariance function*

$$\begin{aligned}
&\Gamma_k\left((\tau_{1,1},\tau_{1,2},....\tau_{1,k}),(\tau_{2,1},\tau_{2,2},....\tau_{2,k})\right)\\
&\quad= E\left[W_k(\tau_{1,1},\tau_{1,2},....\tau_{1,k})\overline{W_k(\tau_{2,1},\tau_{2,2},....\tau_{2,k})}\right]\\
&\quad= E\left[U_t^2\phi_{k,t-1}(\tau_{1,1},\tau_{1,2},....\tau_{1,k})\overline{\phi_{k,t-1}(\tau_{2,1},\tau_{2,2},....\tau_{2,k})}\right]. \qquad (3.16)
\end{aligned}$$

Consequently, by the continuous mapping theorem,

$$B_{k,n}\xrightarrow{\mathrm{d}} B_k=\int_{\Upsilon^k}\left|W_k(\tau_1,\tau_2,....\tau_k)\right|^2\mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k)$$

Note that

$$W_{k,n}(\tau_1,\tau_2,....,\tau_k)=W_{k+1,n}(\tau_1,\tau_2,....,\tau_k,0)$$

and similarly,

$$W_k(\tau_1,\tau_2,....,\tau_k)=W_{k+1}(\tau_1,\tau_2,....,\tau_k,0).$$

Hence

$$\begin{aligned}
&\begin{pmatrix} W_{1,n}(\tau_1)\\ W_{2,n}(\tau_1,\tau_2)\\ \vdots\\ W_{k,n}(\tau_1,\tau_2,\tau_3,....,\tau_k)\end{pmatrix}=\begin{pmatrix} W_{k,n}(\tau_1,0,0,....,0)\\ W_{k,n}(\tau_1,\tau_2,0,....,0)\\ \vdots\\ W_{k,n}(\tau_1,\tau_2,\tau_3,....,\tau_k)\end{pmatrix}\\
&\Rightarrow\begin{pmatrix} W_k(\tau_1,0,0,....,0)\\ W_k(\tau_1,\tau_2,0,....,0)\\ \vdots\\ W_k(\tau_1,\tau_2,\tau_3,....,\tau_k)\end{pmatrix}=\begin{pmatrix} W_1(\tau_1)\\ W_2(\tau_1,\tau_2)\\ \vdots\\ W_k(\tau_1,\tau_2,\tau_3,....,\tau_k)\end{pmatrix},
\end{aligned}$$

where the weak convergence result involved is not hard to verify. It follows now from the continuous mapping theorem that

Lemma 3.6. *Under the conditions of Lemma 3.5,*

$$(B_{1,n},B_{2,n},...,B_{k,n})'\xrightarrow{\mathrm{d}}(B_1,B_2,...,B_k)'.$$

It remains to show that

Lemma 3.7. *Under H_0 and Assumptions 1.1, 2.1, 2.2 and 3.1, and for any positive sequence γ_k satisfying $\sum_{k=1}^{\infty}\gamma_k < \infty$,*

$$\sum_{k=1}^{\infty}\gamma_k B_{k,n} \xrightarrow{\mathrm{d}} \sum_{k=1}^{\infty}\gamma_k B_k.$$

Proof. It follows from Lemma 3.5 that

$$\begin{aligned} E[B_k] &= \int_{\Upsilon^k} E\left[|W_k(\tau_1,\tau_2,....\tau_k)|^2\right]\mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k) \\ &= \int_{\Upsilon^k} E\left[W_k(\tau_1,\tau_2,....\tau_k)\overline{W_k(\tau_1,\tau_2,....\tau_k)}\right] \\ &\quad \times\mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k) \\ &= \int_{\Upsilon^k} E\left[U_t^2\phi_{k,t-1}(\tau_1,\tau_2,....\tau_k)\overline{\phi_{k,t-1}(\tau_1,\tau_2,....\tau_k)}\right] \\ &\quad \times\mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k) \end{aligned}$$

hence, similar to Lemma 3.3 we have

$$\sup_{k\in\mathbb{N}} E[B_k] < \infty. \tag{3.17}$$

Moreover, Lemma 3.6 implies that for each $\ell \in \mathbb{N}$,

$$\sum_{k=1}^{\ell}\gamma_k B_{k,n} \xrightarrow{\mathrm{d}} \sum_{k=1}^{\ell}\gamma_k B_k. \tag{3.18}$$

Let x, $x-\delta$ and $x+\delta$, with $\delta > 0$, be continuity points of the distribution of $\sum_{k=1}^{\infty}\gamma_k B_k$, Then

$$\begin{aligned} &\limsup_{n\to\infty} \Pr\left[\sum_{k=1}^{\infty}\gamma_k B_{k,n} \le x\right] \\ &\quad \le \lim_{n\to\infty}\Pr\left[\sum_{k=1}^{\ell}\gamma_k B_{k,n} \le x\right] = \Pr\left[\sum_{k=1}^{\ell}\gamma_k B_k \le x\right] \\ &\quad = \Pr\left[\sum_{k=1}^{\ell}\gamma_k B_k \le x \text{ and } \sum_{k=\ell+1}^{\infty}\gamma_k B_k \le \delta\right] \end{aligned}$$

$$
\begin{aligned}
&+\Pr\left[\sum_{k=1}^{\ell}\gamma_k B_k \le x \text{ and } \sum_{k=\ell+1}^{\infty}\gamma_k B_k > \delta\right] \\
&\le \Pr\left[\sum_{k=1}^{\infty}\gamma_k B_k \le x+\delta\right] + \Pr\left[\sum_{k=\ell+1}^{\infty}\gamma_k B_k > \delta\right] \\
&\le \Pr\left[\sum_{k=1}^{\infty}\gamma_k B_k \le x+\delta\right] + \delta^{-1}\left(\sum_{k=\ell+1}^{\infty}\gamma_k\right)\sup_{k\in\mathbb{N}} E[B_k],
\end{aligned}
$$

where the first equality follows from (3.18) and the last inequality follows from (3.17) and Chebyshev's inequality for first moments. Hence, letting $\ell \to \infty$ first and then letting $\delta \downarrow 0$ it follows that

$$
\limsup_{n\to\infty} \Pr\left[\sum_{k=1}^{\infty}\gamma_k B_{k,n} \le x\right] \le \Pr\left[\sum_{k=1}^{\infty}\gamma_k B_k \le x\right]. \tag{3.19}
$$

Moreover,

$$
\begin{aligned}
&\Pr\left[\sum_{k=1}^{\infty}\gamma_k B_{k,n} \le x\right] \\
&= \Pr\left[\sum_{k=1}^{\ell}\gamma_k B_{k,n} \le x - \sum_{k=\ell+1}^{\infty}\gamma_k B_{k,n} \text{ and } \sum_{k=\ell+1}^{\infty}\gamma_k B_{k,n} \le \delta\right] \\
&\quad + \Pr\left[\sum_{k=1}^{\ell}\gamma_k B_{k,n} \le x - \sum_{k=\ell+1}^{\infty}\gamma_k B_{k,n} \text{ and } \sum_{k=\ell+1}^{\infty}\gamma_k B_{k,n} > \delta\right] \\
&\ge \Pr\left[\sum_{k=1}^{\ell}\gamma_k B_{k,n} \le x - \sum_{k=\ell+1}^{\infty}\gamma_k B_{k,n} \text{ and } \sum_{k=\ell+1}^{\infty}\gamma_k B_{k,n} \le \delta\right] \\
&\ge \Pr\left[\sum_{k=1}^{\ell}\gamma_k B_{k,n} \le x - \delta \text{ and } \sum_{k=\ell+1}^{\infty}\gamma_k B_{k,n} \le \delta\right] \\
&= \Pr\left[\sum_{k=1}^{\ell}\gamma_k B_{k,n} \le x - \delta\right] \\
&\quad - \Pr\left[\sum_{k=1}^{\ell}\gamma_k B_{k,n} \le x - \delta \text{ and } \sum_{k=\ell+1}^{\infty}\gamma_k B_{k,n} > \delta\right]
\end{aligned}
$$

$$\geq \Pr\left[\sum_{k=1}^{\ell} \gamma_k B_{k,n} \leq x - \delta\right] - \Pr\left[\sum_{k=\ell+1}^{\infty} \gamma_k B_{k,n} > \delta\right]$$

Hence

$$\begin{aligned}
&\liminf_{n\to\infty} \Pr\left[\sum_{k=1}^{\infty} \gamma_k B_{k,n} \leq x\right] \\
&\quad \geq \liminf_{n\to\infty} \Pr\left[\sum_{k=1}^{\ell} \gamma_k B_{k,n} \leq x - \delta\right] - \limsup_{n\to\infty} \Pr\left[\sum_{k=\ell+1}^{\infty} \gamma_k B_{k,n} > \delta\right] \\
&\quad = \Pr\left[\sum_{k=1}^{\ell} \gamma_k B_k \leq x - \delta\right] - \limsup_{n\to\infty} \Pr\left[\sum_{k=\ell+1}^{\infty} \gamma_k B_{k,n} > \delta\right] \\
&\quad \geq \Pr\left[\sum_{k=1}^{\infty} \gamma_k B_k \leq x - \delta\right] - \limsup_{n\to\infty} \Pr\left[\sum_{k=\ell+1}^{\infty} \gamma_k B_{k,n} > \delta\right]
\end{aligned}$$

By Lemma 3.1, (3.15) and Chebyshev's inequality for first moments,

$$\begin{aligned}
&\Pr\left[\sum_{k=\ell+1}^{\infty} \gamma_k B_{k,n} > \delta\right] \\
&\quad \leq \Pr\left[2\sum_{k=\ell+1}^{\infty} \gamma_k \widetilde{B}_{1,k,n} + 2\sum_{k=\ell+1}^{\infty} \gamma_k \widetilde{B}_{2,k,n} > \delta\right] \\
&\quad \leq \Pr\left[\sum_{k=\ell+1}^{\infty} \gamma_k \widetilde{B}_{1,k,n} > \delta/4\right] + \Pr\left[\sum_{k=\ell+1}^{\infty} \gamma_k \widetilde{B}_{2,k,n} > \delta/4\right] \\
&\quad \leq 4.\delta^{-1}\left(\sum_{k=\ell+1}^{\infty} \gamma_k\right) E[U_1^2] + \Pr\left[\sup_{k\in\mathbb{N}} \widetilde{B}_{2,k,n} > \frac{\delta}{4.\left(\sum_{k=\ell+1}^{\infty} \gamma_k\right)}\right],
\end{aligned}$$

hence $\lim_{\ell\to\infty} \limsup_{n\to\infty} \Pr\left[\sum_{k=\ell+1}^{\infty} \gamma_k B_{k,n} > \delta\right] = 0$ and thus

$$\liminf_{n\to\infty} \Pr\left[\sum_{k=1}^{\infty} \gamma_k B_{k,n} \leq x\right] \geq \Pr\left[\sum_{k=1}^{\infty} \gamma_k B_k \leq x - \delta\right].$$

Letting $\delta \downarrow 0$ it follows now that

$$\liminf_{n\to\infty} \Pr\left[\sum_{k=1}^{\infty} \gamma_k B_{k,n} \leq x\right] \geq \Pr\left[\sum_{k=1}^{\infty} \gamma_k B_k \leq x\right] \tag{3.20}$$

Combining (3.19) and (3.20) yields

$$\lim_{n\to\infty} \Pr\left[\sum_{k=1}^{\infty} \gamma_k B_{k,n} \leq x\right] = \Pr\left[\sum_{k=1}^{\infty} \gamma_k B_k \leq x\right],$$

which proves the lemma. ■

Summarizing, the following theorem has been proved.

Theorem 3.1. *Let γ_k be an a priory chosen positive sequence satisfying $\sum_{k=1}^{\infty} \gamma_k < \infty$, and let $L_n < n$ be a subsequence of n satisfying $\lim_{n\to\infty} L_n = \infty$. The test statistic of the WICM test takes the form $\widehat{T}_n = \sum_{k=1}^{L_n} \gamma_k \widehat{B}_{n,k}$, where the ICM test statistics $\widehat{B}_{n,k}$ are defined by* (1.14). *Under H_0 and Assumptions 1.1, 2.1, 2.2 and 3.1,*

$$\widehat{T}_n \xrightarrow{\mathrm{d}} T = \sum_{k=1}^{\infty} \gamma_k B_k, \tag{3.21}$$

where each B_k takes the form

$$B_k = \int_{\Upsilon^k} |W_k(\tau_1, \tau_2, ...\tau_k)|^2 \, \mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k),$$

with $W_k(\tau_1, \tau_2, ...\tau_k)$ a complex-valued zero-mean Gaussian process on Υ^k with covariance function given in (3.16).

4. Consistency of the WICM test

In this section I will show that under H_1 the WICM test statistic $\widehat{T}_n$ in Theorem 3.1 converges in probability to infinity, so that this test is consistent. This will be done in the following three steps.

Lemma 4.1. *Denote*

$$\psi_{k,n}(\tau_1, \tau_2, ...\tau_k) = \frac{1}{n}\sum_{t=1}^{n} U_t \exp\left(\mathbf{i}.\sum_{j=1}^{k} \tau_j' \Phi(Z_{t-j})\right), \tag{4.1}$$

$$\eta_{k,n} = \int_{\Upsilon^k} |\psi_{k,n}(\tau_1, \tau_2, ...\tau_k)|^2 \, \mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k). \tag{4.2}$$

Under H_1 and Assumptions 1.1 and 2.1,

$$\sup_{k\in\mathbb{N}} \sup_{(\tau_1,\tau_2,....\tau_k)\in\Upsilon^k} \left|\widehat{W}_{k,n}(\tau_1,\tau_2,....\tau_k)/\sqrt{n} - \psi_{k,n}(\tau_1,\tau_2,....\tau_k)\right| = o_p(1), \tag{4.3}$$

hence

$$\sup_{k\in\mathbb{N}} \left|\widehat{B}_{k,n}/n - \eta_{k,n}\right| = o_p(1). \tag{4.4}$$

Moreover, for each $k \in \mathbb{N}$,

$$\sup_{(\tau_1,\tau_2,....\tau_k)\in\Upsilon^k} |\psi_{k,n}(\tau_1,\tau_2,....\tau_k) - \psi_k(\tau_1,\tau_2,....\tau_k)| = o_p(1), \tag{4.5}$$

where $\psi_k(\tau_1,\tau_2,....\tau_k)$ is defined by (1.10), *hence*

$$\eta_{k,n} = \eta_k + o_p(1), \tag{4.6}$$

where η_k is defined in (1.15).

Proof. Observe from (3.2) that under H_1,

$$\begin{aligned}
&\widehat{W}_{k,n}(\tau_1,\tau_2,....\tau_k)/\sqrt{n} \\
&\quad = \frac{1}{n}\sum_{t=1}^{n}(Y_t - f_{t-1}(\widehat{\theta}_n))\exp\left(\mathbf{i}.\sum_{j=1}^{k}\tau_j'\Phi(Z_{t-j})\right) \\
&\quad = \frac{1}{n}\sum_{t=1}^{n}U_t\exp\left(\mathbf{i}.\sum_{j=1}^{k}\tau_j'\Phi(Z_{t-j})\right) \\
&\qquad - \frac{1}{n}\sum_{t=1}^{n}\left(f_{t-1}(\widehat{\theta}_n) - f_{t-1}(\theta_0)\right)\exp\left(\mathbf{i}.\sum_{j=1}^{k}\tau_j'\Phi(Z_{t-j})\right),
\end{aligned}$$

where now $U_t = Y_t - f_{t-1}(\theta_0)$ is no longer a martingale difference process. Hence,

$$\begin{aligned}
&\sup_{k\in\mathbb{N}} \sup_{(\tau_1,\tau_2,....\tau_k)\in\Upsilon^k} \left|\widehat{W}_{k,n}(\tau_1,\tau_2,....\tau_k)/\sqrt{n} - \psi_{k,n}(\tau_1,\tau_2,....\tau_k)\right| \\
&\qquad \leq \frac{1}{n}\sum_{t=1}^{n}\left|f_{t-1}(\widehat{\theta}_n) - f_{t-1}(\theta_0)\right|.
\end{aligned}$$

Since by Lemma 2.1,

$$\sup_{\theta\in\Theta}\left|\frac{1}{n}\sum_{t=1}^{n}|f_{t-1}(\theta) - f_{t-1}(\theta_0)| - E\left[|f_{t-1}(\theta) - f_{t-1}(\theta_0)|\right]\right| = o_p(1)$$

and by Lemma 2.1, $\widehat{\theta}_n \xrightarrow{\text{p}} \theta_0$, it follows from the continuity of $E[|f_{t-1}(\theta) - f_{t-1}(\theta_0)|]$ that

$$\begin{aligned} \frac{1}{n}\sum_{t=1}^{n}\left|f_{t-1}(\widehat{\theta}_n) - f_{t-1}(\theta_0)\right| &= E[|f_{t-1}(\theta) - f_{t-1}(\theta_0)|]|_{\theta=\widehat{\theta}_n} + o_p(1) \\ &= o_p(1), \end{aligned}$$

which proves (4.3). The latter implies trivially the result (4.4).

The result (4.5) follows straightforwardly from Lemma 2.1, and (4.6) follows straightforwardly from (4.5). ■

Lemma 4.2. *Under H_1 and Assumptions 1.1 and 2.1,*

$$\sup_{k\in\mathbb{N}} \widehat{B}_{k,n}/n = O_p(1), \ \sup_{k\in\mathbb{N}} \eta_{k,n} = O_p(1), \ \sup_{k\in\mathbb{N}} \eta_k = O(1).$$

Proof. It follows from (4.1) that

$$|\psi_{k,n}(\tau_1, \tau_2, \ldots.\tau_k)|^2 \leq \left(\frac{1}{n}\sum_{t=1}^{n}|U_t|\right)^2 \leq \frac{1}{n}\sum_{t=1}^{n}U_t^2$$

hence by (4.2),

$$\sup_{k\in\mathbb{N}} \eta_{k,n} \leq \frac{1}{n}\sum_{t=1}^{n}U_t^2 = \widehat{Q}_n(\theta_0) = Q(\theta_0) + o_p(1) = E[U_1^2] + o_p(1) \tag{4.7}$$

where the equalities follow from (2.1) and (2.2). Obviously, (4.4) and (4.7) imply that $\sup_{k\in\mathbb{N}} \widehat{B}_{k,n}/n = O_p(1)$.

Similarly, it follows from (1.10) that

$$|\psi_k(\tau_1, \tau_2, \ldots.\tau_k)|^2 \leq (E[|U_1|])^2 \leq E[U_1^2]$$

hence $\sup_{k\in\mathbb{N}} \eta_k \leq E[U_1^2] = O(1)$. ■

Using Lemmas 4.1 and 4.2, I will now prove that the WICM test is consistent.

Theorem 4.1. *Under H_1 and Assumptions 1.1 and 2.1, and with γ_k and L_n as in Theorem 3.1,*

$$\widehat{T}_n/n = \sum_{k=1}^{L_n}\gamma_k\widehat{B}_{k,n}/n \xrightarrow{\text{p}} \sum_{k=1}^{\infty}\gamma_k\eta_k > 0. \tag{4.8}$$

Proof. For fixed $\ell \leq L_n$,

$$\left|\sum_{k=1}^{L_n} \gamma_k \widehat{B}_{k,n}/n - \sum_{k=1}^{\infty} \gamma_k \eta_k\right| \leq \sum_{k=1}^{\ell} \gamma_k |\widehat{B}_{k,n}/n - \eta_k|$$
$$+ \left(\sum_{k=\ell+1}^{\infty} \gamma_k\right)\left(\sup_{k\in\mathbb{N}} \widehat{B}_{k,n}/n + \sup_{k\in\mathbb{N}} \eta_{k,n} + \sup_{k\in\mathbb{N}} \eta_k\right). \tag{4.9}$$

Let $\varepsilon > 0$ and $\delta > 0$ be arbitrary. By Lemma 4.2 we can choose ℓ so large that

$$\limsup_{n\to\infty} \Pr\left[\left(\sum_{k=\ell+1}^{\infty} \gamma_k\right)\left(\sup_{k\in\mathbb{N}} \widehat{B}_{k,n}/n + \sup_{k\in\mathbb{N}} \eta_{k,n} + \sup_{k\in\mathbb{N}} \eta_k\right) > \delta/2\right] < \varepsilon \tag{4.10}$$

and for this ℓ it follows from parts (4.5) and (4.6) of Lemma 4.1 that

$$p\lim_{n\to\infty} \sum_{k=1}^{\ell} \gamma_k |\widehat{B}_{k,n}/n - \eta_k| = 0,$$

hence

$$\lim_{n\to\infty} \Pr\left[\sum_{k=1}^{\ell} \gamma_k |\widehat{B}_{k,n}/n - \eta_k| > \delta/2\right] = 0. \tag{4.11}$$

It follows now from (3.15), (4.9), (4.10) and (4.11) that

$$\limsup_{n\to\infty} \Pr\left[\left|\sum_{k=1}^{L_n} \gamma_k \widehat{B}_{k,n}/n - \sum_{k=1}^{\infty} \gamma_k \eta_k\right| > \delta\right] < \varepsilon,$$

which by the arbitrariness of ε implies that

$$\lim_{n\to\infty} \Pr\left[\left|\sum_{k=1}^{L_n} \gamma_k \widehat{B}_{k,n}/n - \sum_{k=1}^{\infty} \gamma_k \eta_k\right| > \delta\right] = 0.$$

This result and the result of Lemma 1.2 imply (4.8). ■

5. Upper bounds of the critical values

In Theorem 5.1[5] in the addendum to Bierens and Wang (2012) in Chapter 6 it has been shown that for a zero mean complex-valued continuous Gaussian random function W on a compact subset Υ of a Euclidean space,

$$\int_{\Upsilon} |W(\tau)|^2 \mathrm{d}\mu(\tau) = \sum_{j=1}^{\infty} \omega_j \varepsilon_j^2, \tag{5.1}$$

where μ is a probability measure on Υ, the ε_j's are independent standard normally distributed random variables and the ω_j's are positive numbers satisfying

$$\sum_{j=1}^{\infty} \omega_j = \int_{\Upsilon} E[|W(\tau)|^2] \mathrm{d}\mu(\tau) < \infty. \tag{5.2}$$

Moreover, it has been shown in Theorem 5.1 in the addendum to Bierens (1982) in Chapter 2 that similar to Bierens and Ploberger (1997, Theorem 7), (5.1) and (5.2) imply that for all $y > 0$,

$$\Pr\left[\frac{\int_{\Upsilon} |W(\tau)|^2 \mathrm{d}\mu(\tau)}{\int_{\Upsilon} E[|W(\tau)|^2] \mathrm{d}\mu(\tau)} > y\right] \leq \Pr\left[\overline{\chi}_1^2 > y\right], \text{ where } \overline{\chi}_1^2 = \sup_{n \geq 1} \frac{1}{n} \sum_{j=1}^{n} \varepsilon_j^2 \tag{5.3}$$

Clearly, these results apply to each B_k in Theorem 3.1.

I will now show that (5.3) applies to $T = \sum_{k=1}^{\infty} \gamma_k B_k$ in Theorem 3.1 as well.

Theorem 5.1. *Under the conditions of Theorem 3.1, and for fixed $K \in \mathbb{N}$,*

$$T_K = \sum_{k=1}^{K} \gamma_k B_k = \sum_{j=1}^{\infty} \omega_{K,j} \varepsilon_{K,j}^2 \tag{5.4}$$

where the $\omega_{K,j}$'s are non-negative and satisfy $\sum_{j=1}^{\infty} \omega_{K,j} = E[T_K]$, and the $\varepsilon_{K,j}$'s are independent standard normally distributed random variables. Hence,

$$\Pr\left[T_K / E[T_K] > y\right] \leq \Pr\left[\overline{\chi}_1^2 > y\right] \tag{5.5}$$

for all $y > 0$, with $\overline{\chi}_1^2$ the same as in (5.3)*. Moreover, letting $K \to \infty$, it follows that for all $y > 0$,*

$$\Pr\left[T / E[T] > y\right] \leq \Pr\left[\overline{\chi}_1^2 > y\right]. \tag{5.6}$$

[5] This theorem is a corected version of Lemma 4 in Bierens and Wang (2012).

Proof. Recall that $B_k = \int_{\Upsilon^k} |W_k(\tau_1, \tau_2, ...\tau_k)|^2 \mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k)$, where $W_k(.)$ is a continuous complex-valued zero-mean Gaussian process on Υ^k.

Denote

$$W_k^+(\tau_1, \tau_2, ..., \tau_k, \tau_{k+1}) = \sqrt{\gamma_k} W_k(\tau_1, \tau_2, ...\tau_k)\rho_{k+1}(\tau_{k+1}),$$

where $\rho_k(\tau)$ is a sequence of continuous functions on Υ such that for $k, m = 1, 2, 3, ...$,

$$\int_\Upsilon \rho_k(\tau)\rho_m(\tau)\mathrm{d}\mu(\tau) = I(k = m), \quad \int_\Upsilon \rho_k(\tau)\mathrm{d}\mu(\tau) = 0.$$

Such a sequence can always be constructed. For example, let $\{\rho_k(\tau)\}_{k=0}^\infty$ with $\rho_0(\tau) \equiv 1$ be the sequence of orthonormal polynomials on Υ. Then for $1 \leq m < k \leq K$,

$$\begin{aligned}
&\int_{\Upsilon^{K+1}} W_m^+(\tau_1, \tau_2, ..., \tau_{m+1})\overline{W_k^+(\tau_1, \tau_2, ..., \tau_{k+1})}\mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_{K+1}) \\
&= \sqrt{\gamma_m}\sqrt{\gamma_k}\int_{\Upsilon^k} W_m(\tau_1, \tau_2, ..., \tau_m)\rho_{m+1}(\tau_{m+1})\overline{W_k(\tau_1, \tau_2, ..., \tau_k)} \\
&\qquad \times \mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k)\int_\Upsilon \rho_{k+1}(\tau)\mathrm{d}\mu(\tau) \\
&= 0
\end{aligned}$$

whereas for $m = k$,

$$\begin{aligned}
&\int_{\Upsilon^{K+1}} W_k^+(\tau_1, \tau_2, ..., \tau_{k+1})\overline{W_k^+(\tau_1, \tau_2, ..., \tau_{k+1})}\mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_{K+1}) \\
&= \gamma_k \int_{\Upsilon^k} W_k(\tau_1, \tau_2, ...\tau_m)\overline{W_k(\tau_1, \tau_2, ...\tau_k)}\mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k) \\
&\qquad \times \int_\Upsilon \rho_{k+1}(\tau)^2\mathrm{d}\mu(\tau) \\
&= \gamma_k \int_{\Upsilon^k} |W_k(\tau_1, \tau_2, ...\tau_m)|^2\mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k) \\
&= \gamma_k B_k
\end{aligned}$$

Consequently, denoting

$$\overline{W}_K(\tau_1, \tau_2, ..., \tau_{K+1}) = \sum_{k=1}^K W_k^+(\tau_1, \tau_2, ..., \tau_{k+1}),$$

we have

$$\int_{\Upsilon^{K+1}} |\overline{W}_K(\tau_1,\tau_2,...,\tau_{K+1})|^2 \mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_{K+1}) = \sum_{k=1}^{K} \gamma_k B_k. \tag{5.7}$$

Note that for $k < K$, $W_k(\tau_1,\tau_2,....,\tau_k) = W_K(\tau_1,\tau_2,....,\tau_k,0,..,0)$, hence

$$\begin{aligned}\overline{W}_K(\tau_1,\tau_2,...,\tau_{K+1}) &= \sum_{k=1}^{K-1} \sqrt{\gamma_k} W_K(\tau_1,\tau_2,....,\tau_k,0,..,0)\rho_{k+1}(\tau_{k+1}) \\ &\quad + \sqrt{\gamma_K} W_K(\tau_1,\tau_2,....,\tau_K)\rho_{K+1}(\tau_{K+1})\end{aligned}$$

is a continuous complex-valued zero mean Gaussian process on Υ^{K+1} because $W_K(\tau_1,\tau_2,....,\tau_K)$ is a continuous complex-valued zero mean Gaussian process on Υ^K and the functions $\rho_k(\tau)$ are continuous and nonrandom. Therefore, similar to (5.1), $T_K = \sum_{k=1}^{K} \gamma_k B_k$ has the representation (5.4), so that similar to (5.2),

$$\Pr\left[\frac{T_K}{E[T_K]} > y\right] \leq \Pr\left[\sup_{n\in\mathbb{N}} \frac{1}{n}\sum_{j=1}^{n} \varepsilon_{K,j}^2 > y\right] = \Pr\left[\overline{\chi}_1^2 > y\right].$$

Finally, recall from (3.17) that $\sup_{k\in\mathbb{N}} E[B_k] < \infty$, which implies that

$$p\lim_{K\to\infty} T_K = T \text{ and } \lim_{K\to\infty} E[T_K] = E[T]$$

and thus $T_K/E[T_K] \xrightarrow{\mathrm{d}} T/E[T]$ as $K \to \infty$. The latter is equivalent to

$$\lim_{K\to\infty} \Pr\left[T_K/E[T_K] \leq y\right] = \Pr\left[T/E[T] \leq y\right] \tag{5.8}$$

in the continuity points y of the distribution function of $T/E[T]$.

However, in Theorem 5.3 at the end of this section it will be shown that the distribution function of T is continuous on $(0,\infty)$, and the same applies to $T/E[T]$. Consequently, equality (5.8) holds for all $y > 0$. ■

To make these result operational for the WICM test statistic $\widehat{T}_n = \sum_{k=1}^{L_n} \gamma_k \widehat{B}_{n,k}$ we need estimates $\widehat{\Gamma}_{n,k}$ of the covariance functions Γ_k of the $W_k(\tau_1,\tau_2,....,\tau_k)$'s such that

$$\begin{aligned}&\widehat{E}(\widehat{T}_n) \stackrel{\text{def.}}{=} \sum_{k=1}^{L_n} \gamma_k \int_{\Upsilon^k} \widehat{\Gamma}_{n,k}\left((\tau_1,\tau_2,....,\tau_k),(\tau_1,\tau_2,....,\tau_k)\right) \mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k) \\ &\xrightarrow{\mathrm{p}} \sum_{k=1}^{\infty} \gamma_k \int_{\Upsilon^k} \Gamma_k\left((\tau_1,\tau_2,....,\tau_k),(\tau_1,\tau_2,....,\tau_k)\right) \mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k) = E[T].\end{aligned}$$

In Chapter 2 I have shown how to derive closed form expressions of the $\widehat{B}_{n,k}$'s and the consistent estimates $\widehat{\Gamma}_{n,k}$ of Γ_k for the case that Υ is a hypercube centered around the zero vector and μ is the uniform probability measure on Υ. The formulas involved can easily adapted to the present case, which is left to the reader.

Now the following results hold for the standardized WICM test $\widehat{T}_n/\widehat{E}(\widehat{T}_n)$.

Theorem 5.2. *For* $\alpha \in (0,1)$, *let* $\overline{c}(\alpha)$ *be such that* $\Pr[\overline{\chi}_1^2 > \overline{c}(\alpha)] = \alpha$, *and let* $\widehat{E}(\widehat{T}_n)$ *be a consistent estimator of* $E[T]$ *under* H_0, *and* $\widehat{E}(\widehat{T}_n) = O_p(1)$ *under* H_1. *Then under the conditions of Theorems 3.1 and 4.1, respectively,*

$$\begin{aligned}\lim_{n\to\infty} \Pr\left[\widehat{T}_n/\widehat{E}(\widehat{T}_n) > \overline{c}(\alpha)\right] &\leq \alpha \text{ under } H_0,\\ \lim_{n\to\infty} \Pr\left[\widehat{T}_n/\widehat{E}(\widehat{T}_n) > \overline{c}(\alpha)\right] &= 1 \text{ under } H_1.\end{aligned}$$

Remarks.

- The condition that $\widehat{E}(\widehat{T}_n) = O_p(1)$ under H_1 is not restrictive, because without loss of generality we may base $\widehat{E}(\widehat{T}_n)$ on a bootstrap version of $\widehat{T}_n$ for which $\widehat{E}(\widehat{T}_n) = O_p(1)$ under H_1. See (6.2) in the next section.
- The values of $\overline{c}(\alpha)$ for $\alpha = 0.01$, $\alpha = 0.05$ and $\alpha = 0.1$ have been calculated in Bierens and Ploberger (1997), i.e.,

$$\overline{c}(0.01) = 6.81,\ \overline{c}(0.05) = 4.26,\ \overline{c}(0.10) = 3.23. \tag{5.9}$$

Finally, I will show that the limiting distribution of the WICM test is continuous.

Theorem 5.3. *Let* $G(x) = \Pr[T \leq x]$, *where* T *is the asymptotic null distribution of the WICM test in Theorem 3.1. Under the conditions of Theorem 3.1,* $G(x)$ *is continuous on* $(0,\infty)$.

Proof. It follows from (5.7) and Theorem 6.5 in the addendum to Bierens (1982) in Chapter 2 that the distribution function

$$G_K(x) = \Pr\left[\sum_{k=1}^{K}\gamma_k B_k \leq x\right]$$

is continuous on $(0,\infty)$. Moreover, note that

$$\sum_{k=1}^{K}\gamma_k B_k \xrightarrow{\mathrm{d}} \sum_{k=1}^{\infty}\gamma_k B_k$$

as $K\to\infty$, hence

$$\lim_{K\to\infty} G_K(x) = G(x)$$

in the continuity points of G.

To prove that $G(x)$ is continuous on $(0,\infty)$, suppose that $x_0>0$ is a discontinuity point of G, i.e.,

$$G(x_0) - \lim_{\delta\downarrow 0} G(x_0-\delta) = \varepsilon > 0.$$

Hence, for all $\delta>0$ such that $x_0-\delta$ is a continuity point of G,

$$\varepsilon \le G(x_0) - G(x_0-\delta) = G(x_0) - \lim_{K\to\infty} G_K(x_0-\delta) \tag{5.10}$$

Since G_K is continuous, for each K there exists a $\delta_K>0$ such that $G_K(x_0) - G_K(x_0-\delta_K) < \varepsilon/2$, and since for $\delta\in(0,\delta_K]$, $G_K(x_0)-G_K(x_0-\delta)<\varepsilon/2$ as well, we may without loss of generality assume that $\lim_{K\to\infty}\delta_K = 0$. Then for arbitrary $\delta>0$ and K so large that $\delta>\delta_K$, $G_K(x_0-\delta_K)\ge G_K(x_0-\delta)$, hence

$$\begin{aligned}
G(x_0) - G_K(x_0-\delta) &\le G(x_0) - G_K(x_0-\delta_K)\\
&= G(x_0) - G_K(x_0) + G_K(x_0) - G_K(x_0-\delta_K)\\
&\le G(x_0) - G_K(x_0) + \varepsilon/2\\
&\le \varepsilon/2
\end{aligned}$$

where that last inequality follows from $G(x_0)\le G_K(x_0)$. Again, assuming that $x_0-\delta$ is a continuity point of G, we have

$$G(x_0) - G(x_0-\delta) = G(x_0) - \lim_{K\to\infty} G_K(x_0-\delta) \le \varepsilon/2. \tag{5.11}$$

However, (5.10) and (5.11) contradict, which proves that $G(x)$ does not have any discontinuity points. ■

6. Bootstrap

In this section I will extend the bootstrap procedure in the addendum to Bierens (1982) in Chapter 2 to the present case. For that purpose it is necessary to strengthen some of the conditions in Assumptions 2.2 and the condition in Assumption 3.1 a little bit, as follows.

Assumption 6.1. *Let θ_0 be defined by* (1.3). *The following conditions hold, regardless whether H_0 is true or not.*
(a) Θ is convex and θ_0 is an interior point of Θ.
(b) $f_{t-1}(\theta)$ is a.s. twice continuously differentiable in the components $\theta_1, \theta_2, ..., \theta_m$ of $\theta \in \Theta$;
(c) $E\left[\sup_{\theta\in\Theta} ||\nabla f_{t-1}(\theta)||^2\right] < \infty$.
(d) $E\left[\sup_{\theta\in\Theta}(Y_t - f_{t-1}(\theta))^2.||\nabla f_{t-1}(\theta)||^2\right] < \infty$.
(e) $E\left[\sup_{\theta\in\Theta}(Y_t - f_{t-1}(\theta))^2||\nabla f_{t-1}(\theta)||\right] < \infty$.[6]
(f) For $i_1, i_2 = 1, 2, ..., m$, $E[\sup_{\theta\in\Theta} |Y_t - f_{t-1}(\theta)|.|(\partial/\partial\theta_{i_1})(\partial/\partial\theta_{i_2})f_{t-1}(\theta)|] < \infty$.
(g) There exists a sequence M_t of random variables such that for all $\theta_1, \theta_2 \in \Theta$,

$$||(Y_t - f_{t-1}(\theta_1))\nabla f_{t-1}(\theta_1) - (Y_t - f_{t-1}(\theta_2))\nabla f_{t-1}(\theta_2)|| \leq M_t.||\theta_1 - \theta_2||$$

and

$$\limsup_{n\to\infty} \frac{1}{n}\sum_{t=1}^{n} E[M_t^2] < \infty.$$

(h) For all $\theta \in \Theta$ the matrix $A_2(\theta) = E[(\nabla f_{t-1}(\theta))(\nabla f_{t-1}(\theta))']$ is nonsingular.

Except for part (g), this assumption encompasses the conditions in Assumptions 2.2 and 3.1, for ease of reference. Using the mean value theorem, Assumption 6.1(g) can be broken down in more primitive conditions. In particular, denoting $\nabla^2 f_{t-1}(\theta) = (\partial/\partial\theta)(\partial/\partial\theta')f_{t-1}(\theta)$, and defining the matrix norm as $||A|| = \sqrt{\text{trace}(AA')}$, it follows from the mean value theorem that

$$\begin{aligned}
&||(Y_t - f_{t-1}(\theta_1))\nabla f_{t-1}(\theta_1) - (Y_t - f_{t-1}(\theta_2))\nabla f_{t-1}(\theta_2)|| \\
&\leq |Y_t - f_{t-1}(\theta_1)|.||\nabla f_{t-1}(\theta_1) - \nabla f_{t-1}(\theta_2)|| + |f_{t-1}(\theta_1) - f_{t-1}(\theta_2)|.||\nabla f_{t-1}(\theta_2)|| \\
&\leq \left(\sup_{\theta\in\Theta}|Y_t - f_{t-1}(\theta)|.\sup_{\theta\in\Theta}||\nabla^2 f_{t-1}(\theta)|| + \sup_{\theta\in\Theta}||\nabla f_{t-1}(\theta)||^2\right)||\theta_1 - \theta_2||.
\end{aligned}$$

[6] C.f. Assumption 3.1.

Theorem 6.1. *Denote*

$$\begin{aligned} W^*_{i,k,n}(\tau_1,\tau_2,....\tau_k) &= \frac{1}{\sqrt{n}}\sum_{t=1}^{n}\varepsilon_{i,t}(Y_t - f_{t-1}(\theta_0))\phi_{k,t-1}(\tau_1,\tau_2,....\tau_k), \\ B^*_{i,k,n} &= \int_{\Upsilon^k}\left|W^*_{i,k,n}(\tau_1,\tau_2,....\tau_k)\right|^2 \mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k), \\ T^*_{i,n} &= \sum_{k=1}^{L_n}\gamma_k B^*_{i,k,n},\ i=1,2,...,M, \end{aligned}$$

where the $\varepsilon_{i,t}$*'s are independent random drawings from the standard normal distribution,* M *is the number of bootstraps,* $\phi_{k,t-1}(\tau_1,\tau_2,....\tau_k)$ *is defined by* (3.5), *and* γ_k *and* L_n *are the same as in Theorem 3.1. Under Assumptions 1.1, 2.1 and 6.1 the following results hold.*
(a) For each $k\in\mathbb{N}$ *and* $i=1,2,...,M$,

$$W^*_{i,k,n}(\tau_1,\tau_2,....\tau_k) \Rightarrow W^*_{i,k}(\tau_1,\tau_2,....\tau_k)$$

on Υ^k, *where* $W^*_{i,k}(\tau_1,\tau_2,....\tau_k)$ *is a zero-mean complex-valued Gaussian process on* Υ^k *with covariance function*

$$\begin{aligned} &\Gamma^*_k\left((\tau_{1,1},\tau_{1,2},....\tau_{1,k}),(\tau_{2,1},\tau_{2,2},....\tau_{2,k})\right) \\ &\quad = E\left[W^*_{i.k}(\tau_{1,1},\tau_{1,2},....\tau_{1,k})\overline{W^*_{i.k}(\tau_{2,1},\tau_{2,2},....\tau_{2,k})}\right] \\ &\quad = E\left[(Y_t - f_{t-1}(\theta_0))^2\phi_{k,t-1}(\tau_{1,1},\tau_{1,2},....\tau_{1,k})\overline{\phi_{k,t-1}(\tau_{2,1},\tau_{2,2},....\tau_{2,k})}\right] \\ &\quad = E\left[\left(Y_t - E\left[Y_t|\mathcal{F}^{t-1}_{-\infty}\right]\right)^2\phi_{k,t-1}(\tau_{1,1},\tau_{1,2},....\tau_{1,k})\overline{\phi_{k,t-1}(\tau_{2,1},\tau_{2,2},....\tau_{2,k})}\right] \\ &\qquad + E\Bigg[\left(f_{t-1}(\theta_0) - E\left[Y_t|\mathcal{F}^{t-1}_{-\infty}\right]\right)^2 \\ &\qquad\qquad \times\phi_{k,t-1}(\tau_{1,1},\tau_{1,2},....\tau_{1,k})\overline{\phi_{k,t-1}(\tau_{2,1},\tau_{2,2},....\tau_{2,k})}\Bigg] \\ &\quad = \Gamma_k\left((\tau_{1,1},\tau_{1,2},....\tau_{1,k}),(\tau_{2,1},\tau_{2,2},....\tau_{2,k})\right) \\ &\qquad + E\Bigg[\left(f_{t-1}(\theta_0) - E\left[Y_t|\mathcal{F}^{t-1}_{-\infty}\right]\right)^2 \\ &\qquad\qquad \times\phi_{k,t-1}(\tau_{1,1},\tau_{1,2},....\tau_{1,k})\overline{\phi_{k,t-1}(\tau_{2,1},\tau_{2,2},....\tau_{2,k})}\Bigg] \qquad (6.1) \end{aligned}$$

where Γ_k *is the covariance function* (3.16). *Hence,* $B^*_{i,k,n} \xrightarrow{\text{d}} B^*_{i,k}$, *and*

$$T_{i,n} \xrightarrow{\text{d}} T^*_i = \sum_{k=1}^{\infty} \gamma_k B^*_{i,k}$$

where

$$B^*_{i,k} = \int_{\Upsilon^k} \left| W^*_{i,k}(\tau_1, \tau_2, ...\tau_k) \right|^2 \text{d}\mu(\tau_1)\text{d}\mu(\tau_2)...\text{d}\mu(\tau_k).$$

(*b*) *Moreover,*

$$(T^*_{1,n}, T^*_{2,n}, ..., T^*_{M,n})' \xrightarrow{\text{d}} (T^*_1, T^*_2, ..., T^*_M)',$$

where $T^*_1, T^*_2, ..., T^*_M$ *are i.i.d.. Furthermore, under* H_0,

$$(T^*_{1,n}, T^*_{2,n}, ..., T^*_{M,n})' \xrightarrow{\text{d}} (T_1, T_2, ..., T_M)',$$

where $T_1, T_2, ..., T_M$ *are i.i.d. as* T *in* (3.21) *and are independent of* T.

Proof. Part (a) follows similar to Lemma 3.5 and Theorem 3.1. As to part (b), it follows from the independence of $\varepsilon_{i,t}$'s that, conditional on the data, $T_1, T_2, ..., T_M$ are i.i.d. and are therefore i.i.d. unconditionally as well, as is not hard to verify.[7] The same applies to $T^*_1, T^*_2, ..., T^*_M$. Moreover, under H_0 the covariance function (6.1) is exactly the same as (3.16), hence $T_i \sim T$. Furthermore, since the $\varepsilon_{i,t}$'s are drawn independently of the data, $(T_1, T_2, ..., T_M)'$ is independent of T. ■

If it were possible to compute the $T^*_{i,n}$'s then the asymptotic $\alpha \times 100\%$ critical values $c(\alpha)$ of the WICM test can be approximated by sorting the $T^*_{i,n}$'s in decreasing order and then use $\overline{c}_{n,M}(\alpha) = T^*_{[\alpha M],n}$ as the approximation of $c(\alpha)$, where $[\alpha M]$ denotes the largest natural number less or equal to αM.

[7] Show first that the characteristic function of each T_i conditional on the data equals the unconditional characteristic function $E[\exp(\mathbf{i}\xi T_i)]$, and that these characteristic functions are all equal. Then, with $\mathcal{F}^{\infty}_{-\infty}$ the σ-algebra generated by $\{Z_t\}^{\infty}_{t=-\infty}$, we have

$$\begin{aligned} E\left[\exp\left(\mathbf{i}.\sum_{j=1}^{M}\xi_j T_j\right)\right] &= E\left(E\left[\exp\left(\mathbf{i}.\sum_{j=1}^{M}\xi_j T_j\right)\middle| \mathcal{F}^{\infty}_{-\infty}\right]\right) \\ &= E\left(\prod_{j=1}^{M} E\left[\exp\left(\mathbf{i}.\sum_{j=1}^{M}\xi_j T_j\right)\middle| \mathcal{F}^{\infty}_{-\infty}\right]\right) = \prod_{j=1}^{M} E\left[\exp\left(\mathbf{i}.\sum_{j=1}^{M}\xi_j T_j\right)\right]. \end{aligned}$$

However, obviously the computation of the $T^*_{i,n}$'s is not possible. Therefore, to make this bootstrap procedure feasible, we need to construct feasible bootstrap WICM statistics $\widetilde{T}_{i,n}$ such that $\widetilde{T}_{i,n} = T^*_{i,n} + o_p(1)$, as follows.

Denote for $i = 1, 2, ..., M$,

$$\begin{aligned}
\widehat{b}_{n,k}(\tau_1, \tau_2, ...\tau_k|\theta) &= \frac{1}{n}\sum_{t=1}^{n} \nabla f_{t-1}(\theta) \exp\left(\mathbf{i}\sum_{j=1}^{k} \tau_j' \Phi(Z_{t-j})\right), \\
\widehat{A}_{2,n}(\theta) &= \frac{1}{n}\sum_{t=1}^{n} \nabla f_{t-1}(\theta)(\nabla f_{t-1}(\theta))' \\
\widehat{\phi}_{k,n,t-1}(\tau_1, \tau_2, ...\tau_k|\theta) &= \exp\left(\mathbf{i}\sum_{j=1}^{k} \tau_j' \Phi(Z_{t-j})\right) \\
&\quad -\widehat{b}_{n,k}(\tau_1, \tau_2, ...\tau_k|\theta)' \widehat{A}_{2,n}^{-1}(\theta) \nabla f_{t-1}(\theta), \\
\widetilde{W}_{i,k,n}(\tau_1, \tau_2, ...\tau_k|\theta) &= \frac{1}{\sqrt{n}}\sum_{t=1}^{n} \varepsilon_{i,t}(Y_t - f_{t-1}(\theta))\widehat{\phi}_{k,n,t-1}(\tau_1, \tau_2, ...\tau_k|\theta), \\
\widetilde{B}_{i,k,n} &= \int_{\Upsilon^k} \left|\widetilde{W}_{i,k,n}(\tau_1, \tau_2, ...\tau_k|\widehat{\theta}_n)\right|^2 \mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k), \\
\widetilde{T}_{i,n} &= \sum_{k=1}^{L_n} \gamma_k \widetilde{B}_{i,k,n}, \qquad (6.2)
\end{aligned}$$

where the $\varepsilon_{i,t}$'s are the same as in Theorem 6.1.

We can write

$$\begin{aligned}
&\widetilde{W}_{i,k,n}(\tau_1, \tau_2, ...\tau_k|\widehat{\theta}_n) \\
&= \frac{1}{\sqrt{n}}\sum_{t=1}^{n} \varepsilon_{i,t}(Y_t - f_{t-1}(\theta_0)) \exp\left(\mathbf{i}\sum_{j=1}^{k} \tau_j' \Phi(Z_{t-j})\right) \\
&\quad - \widehat{a}_{n,k,i}(\widehat{\theta}_n, \tau_1, \tau_2, ...\tau_k) \\
&\quad - \widehat{b}_{n,k}(\tau_1, \tau_2, ...\tau_k|\widehat{\theta}_n)' \widehat{A}_{2,n}^{-1}(\widehat{\theta}_n)\widehat{d}_{n,k,i}(\widehat{\theta}_n, \tau_1, \tau_2, ...\tau_k)
\end{aligned}$$

where

$$\widehat{a}_{n,k,i}(\theta, \tau_1, \tau_2, ...\tau_k) = \frac{1}{\sqrt{n}}\sum_{t=1}^{n} \varepsilon_{i,t}\left(f_{t-1}(\theta) - f_{t-1}(\theta_0)\right)$$

$$
\begin{aligned}
& \qquad \times \exp\left(\mathbf{i}\sum_{j=1}^{k}\tau_j'\Phi(Z_{t-j})\right), \\
\widehat{d}_{n,k,i}(\theta,\tau_1,\tau_2,....\tau_k) &= \frac{1}{\sqrt{n}}\sum_{t=1}^{n}\varepsilon_{i,t}\Psi_{k,t}(\theta,\tau_1,\tau_2,....\tau_k), \text{ with} \\
\Psi_{k,t}(\theta,\tau_1,\tau_2,....\tau_k) &= (Y_t - f_{t-1}(\theta))\nabla f_{t-1}(\theta)\exp\left(\mathbf{i}\sum_{j=1}^{k}\tau_j'\Phi(Z_{t-j})\right).
\end{aligned}
$$

Similarly,

$$
\begin{aligned}
& W^*_{i,k,n}(\tau_1,\tau_2,....\tau_k) \\
& \quad = \frac{1}{\sqrt{n}}\sum_{t=1}^{n}\varepsilon_{i,t}(Y_t - f_{t-1}(\theta_0))\exp\left(\mathbf{i}\sum_{j=1}^{k}\tau_j'\Phi(Z_{t-j})\right) \\
& \qquad - b_k(\tau_1,\tau_2,....\tau_k)'A_2^{-1}\widehat{d}_{n,k,i}(\theta_0,\tau_1,\tau_2,....\tau_k).
\end{aligned}
$$

I will show first that

$$
\sup_{(\tau_1,\tau_2,....\tau_k)\in\Upsilon^k}\left|\widetilde{W}_{i,k,n}(\tau_1,\tau_2,....\tau_k|\widehat{\theta}_n) - W^*_{i,k,n}(\tau_1,\tau_2,....\tau_k)\right| = o_p(1), \tag{6.3}
$$

using the following three lemmas, and then that $\widetilde{T}_{i,n} = T^*_{i,n} + o_p(1)$.

Lemma 6.1. *Under the conditions of Theorem 6.1 and for each k and i, $\widehat{d}_{n,k,i}(\theta,\tau_1,\tau_2,....\tau_k)$ is tight on $\Theta\times\Upsilon^k$. Consequently,*

$$
\sup_{(\tau_1,\tau_2,....\tau_k)\in\Upsilon^k}\left\|\widehat{d}_{n,k,i}(\widehat{\theta}_n,\tau_1,\tau_2,....\tau_k) - \widehat{d}_{n,k,i}(\theta_0,\tau_1,\tau_2,....\tau_k)\right\| = o_p(1), \tag{6.4}
$$

where here and in the sequel the norm $||.||$ on $\mathbb{C}^m$ is defined as $||a+\mathbf{i}.b|| = \sqrt{a'a+b'b}$.

Proof. According to Lemma A.1 in Bierens and Ploberger (1997), in multivariate form, $\mathrm{Re}[\widehat{d}_{n,k,i}(\theta,\tau_1,\tau_2,....\tau_k)]$ is tight on $\Theta\times\Upsilon^k$ if for any pair

$$
(\theta_1,\tau_{1,1},\tau_{1,2},....\tau_{1,k}),\ (\theta_2,\tau_{2,1},\tau_{2,2},....\tau_{2,k}) \in \Theta\times\Upsilon^k,
$$

the following Lipschitz condition holds:

$$
\begin{aligned}
&|| \operatorname{Re}[\Psi_{k,t}(\theta_1, \tau_{1,1}, \tau_{1,2}, ...\tau_{1,k})] - \operatorname{Re}[\Psi_{k,t}(\theta_2, \tau_{2,1}, \tau_{2,2}, ...\tau_{2,k})]|| \\
&\qquad \leq K_{k,t}\sqrt{||\theta_1 - \theta_2||^2 + \sum_{j=1}^{k} ||\tau_{2,j} - \tau_{1,j}||^2}, \qquad (6.5)
\end{aligned}
$$

where $K_{k,t}$ is a sequence of random variables such that

$$
\limsup_{n\to\infty} \frac{1}{n}\sum_{t=1}^{n} E\left[\varepsilon_{i,t}^2 K_{k,t}^2\right] < \infty, \qquad (6.6)
$$

and for an arbitrary point $(\theta, \tau_1, \tau_2, ...\tau_k) \in \Theta \times \Upsilon^k$,

$$
\limsup_{n\to\infty} \frac{1}{n}\sum_{t=1}^{n} E\left[\varepsilon_{i,t}^2 || \operatorname{Re}(\Psi_{k,t}(\theta, \tau_1, \tau_2, ...\tau_k))||^2\right] < \infty. \qquad (6.7)
$$

Condition (6.5) follows from

$$
\begin{aligned}
&|| \operatorname{Re}[\Psi_{k,t}(\theta_1, \tau_{1,1}, \tau_{1,2}, ...\tau_{1,k})] - \operatorname{Re}[\Psi_{k,t}(\theta_2, \tau_{2,1}, \tau_{2,2}, ...\tau_{2,k})]|| \\
&= \left\| (Y_t - f_{t-1}(\theta_1))\nabla f_{t-1}(\theta_1)\cos\left(\sum_{j=1}^{k}\tau_{1,j}'\Phi(Z_{t-j})\right)\right. \\
&\qquad \left. - (Y_t - f_{t-1}(\theta_2))\nabla f_{t-1}(\theta_2)\cos\left(\sum_{j=1}^{k}\tau_{2,j}'\Phi(Z_{t-j})\right)\right\| \\
&\leq ||(Y_t - f_{t-1}(\theta_1))\nabla f_{t-1}(\theta_1) - (Y_t - f_{t-1}(\theta_2))\nabla f_{t-1}(\theta_2)|| \\
&\quad + \left\|(Y_t - f_{t-1}(\theta_2))\nabla f_{t-1}(\theta_2)\sum_{j=1}^{k}(\tau_{2,j} - \tau_{1,j})'\Phi(Z_{t-j})\right\| \\
&\leq M_{t-1}.||\theta_1 - \theta_2|| \\
&\quad + \left(\sup_z ||\Phi(z)||\right)\sup_{\theta\in\Theta}\left(|Y_t - f_{t-1}(\theta)|.||\nabla f_{t-1}(\theta)||\right)\left(\sum_{j=1}^{k}||\tau_{2,j} - \tau_{1,j}||\right) \\
&\leq K_{k,t}\sqrt{||\theta_1 - \theta_2||^2 + \sum_{j=1}^{k}||\tau_{2,j} - \tau_{1,j}||^2},
\end{aligned}
$$

where the second inequality follows from Assumption 6.1(g), and the last inequality from

$$K_{k,t} = \sqrt{k+1}\left(M_t + \left(\sup_z ||\Phi(z)||\right)\sup_{\theta\in\Theta}\left(|Y_t - f_{t-1}(\theta)|.||\nabla f_{t-1}(\theta)||\right)\right)$$

and

$$||\theta_1 - \theta_2|| + \sum_{j=1}^{k} ||\tau_{2,j} - \tau_{1,j}|| \leq \sqrt{k+1}\sqrt{||\theta_1 - \theta_2||^2 + \sum_{j=1}^{k} ||\tau_{2,j} - \tau_{1,j}||^2}.$$

Moreover, condition (6.6) follows from

$$\begin{aligned}
&\limsup_{n\to\infty} \frac{1}{n}\sum_{t=1}^{n} E\left[\varepsilon_{i,t}^2 K_{k,t}^2\right] = \limsup_{n\to\infty} \frac{1}{n}\sum_{t=1}^{n} E\left[K_{k,t}^2\right] \\
&\leq 2(k+1)\limsup_{n\to\infty} \frac{1}{n}\sum_{t=1}^{n} E[M_t^2] \\
&+ 2(k+1)\sup_z ||\Phi(z)||^2 E\left[\sup_{\theta\in\Theta}(Y_t - f_{t-1}(\theta))^2.||\nabla f_{t-1}(\theta)||^2\right] \\
&< \infty,
\end{aligned}$$

where the last inequality is due to parts (d) and (g) of Assumption 6.1.

Finally, condition (6.7) follows from part (d) of Assumption 6.1 and the easy inequality

$$E\left[||\operatorname{Re}(\Psi_{k,t}(\theta,\tau_1,\tau_2,....\tau_k))||^2\right] \leq E\left[\sup_{\theta\in\Theta}(Y_t - f_{t-1}(\theta))^2||\nabla f_{t-1}(\theta)||^2\right].$$

Thus, $\operatorname{Re}[\widehat{d}_{n,k,i}(\theta,\tau_1,\tau_2,....\tau_k)]$ is tight on $\Theta\times\Upsilon^k$. Obviously, the same holds for $\operatorname{Im}[\widehat{d}_{n,k,i}(\theta,\tau_1,\tau_2,....\tau_k)]$ and thus for $\widehat{d}_{n,k,i}(\theta,\tau_1,\tau_2,....\tau_k)$ as well.

Using the martingale difference central limit theorem of McLeish (1974) it follows straightforwardly that the finite distributions of $\widehat{d}_{n,k,i}(\theta,\tau_1,\tau_2,....\tau_k)$ converge to multivariate zero-mean complex-valued normal distributions, hence

$$\widehat{d}_{n,k,i}(\theta,\tau_1,\tau_2,....\tau_k) \Rightarrow d_{k,i}(\theta,\tau_1,\tau_2,....\tau_k)$$

on $\Theta \times \Upsilon^k$, where $d_{k,i}(\theta, \tau_1, \tau_2, ...\tau_k)$ is a an m-dimensional zero-mean complex-valued Gaussian process on $\Theta \times \Upsilon^k$. Similarly,

$$\begin{aligned}
&\widehat{d}_{n,k,i}(\theta, \tau_1, \tau_2, ...\tau_k) - \widehat{d}_{n,k,i}(\theta_0, \tau_1, \tau_2, ...\tau_k) \\
&\quad \Rightarrow d_{k,i}(\theta, \tau_1, \tau_2, ...\tau_k) - d_{k,i}(\theta_0, \tau_1, \tau_2, ...\tau_k)
\end{aligned} \tag{6.8}$$

on $\Theta \times \Upsilon^k$.

Now let $\Theta_0(\varepsilon) = \{\theta \in \mathbb{R}^m : ||\theta - \theta_0|| \leq \varepsilon\}$, where $\varepsilon > 0$ is so small that $\Theta_0(\varepsilon) \subset \Theta$. and denote

$$\begin{aligned}
\widehat{\Delta}_{n,k,i}(\theta) &= \sup_{(\tau_1,\tau_2,....\tau_k)\in\Upsilon^k} \left\| \widehat{d}_{n,k,i}(\theta, \tau_1, \tau_2, ...\tau_k) - \widehat{d}_{n,k,i}(\theta_0, \tau_1, \tau_2, ...\tau_k) \right\|, \\
\Delta_{k,i}(\theta) &= \sup_{(\tau_1,\tau_2,....\tau_k)\in\Upsilon^k} \left\| d_{k,i}(\theta, \tau_1, \tau_2, ...\tau_k) - d_{k,i}(\theta_0, \tau_1, \tau_2, ...\tau_k) \right\|.
\end{aligned}$$

Then by the continuous mapping theorem, (6.8) implies

$$\sup_{\theta\in\Theta_0(\varepsilon)} \widehat{\Delta}_{n,k,i}(\theta) \xrightarrow{\mathrm{d}} \sup_{\theta\in\Theta_0(\varepsilon)} \Delta_{k,i}(\theta)$$

and thus for any $\delta > 0$,

$$\begin{aligned}
\Pr\left[\widehat{\Delta}_{n,k,i}(\widehat{\theta}_n) > \delta\right] &= \Pr\left[\widehat{\Delta}_{n,k,i}(\widehat{\theta}_n) > \delta \text{ and } \widehat{\theta}_n \in \Theta_0(\varepsilon)\right] \\
&\quad + \Pr\left[\widehat{\Delta}_{n,k,i}(\widehat{\theta}_n) > \delta \text{ and } \widehat{\theta}_n \notin \Theta_0(\varepsilon)\right] \\
&\leq \Pr\left[\sup_{\theta\in\Theta_0(\varepsilon)} \widehat{\Delta}_{n,k,i}(\theta) > \delta\right] + \Pr\left[||\widehat{\theta}_n - \theta_0|| > \varepsilon\right] \\
&\to \Pr\left[\sup_{\theta\in\Theta_0(\varepsilon)} \Delta_{k,i}(\theta) > \delta\right]
\end{aligned}$$

Since $\lim_{\varepsilon\downarrow 0} \sup_{\theta\in\Theta_0(\varepsilon)} \Delta_{k,i}(\theta) = \Delta_{k,i}(\theta_0) = 0$ a.s. and thus

$$\lim_{\varepsilon\downarrow 0} \Pr\left[\sup_{\theta\in\Theta_0(\varepsilon)} \Delta_{k,i}(\theta) > \delta\right] = 0,$$

it follows now that for any $\delta > 0$, $\lim_{n\to\infty} \Pr\left[\widehat{\Delta}_{n,k,i}(\widehat{\theta}_n) > \delta\right] = 0$. The latter proves (6.4). ■

Similarly,

Lemma 6.2. *Under the conditions of Theorem 6.1 and for each k and i, $\widehat{a}_{n,k,i}(\theta,\tau_1,\tau_2,....\tau_k)$ is tight on $\Theta\times\Upsilon^k$. Consequently,*

$$\sup_{(\tau_1,\tau_2,....\tau_k)\in\Upsilon^k}\left|\widehat{a}_{n,k,i}(\widehat{\theta}_n,\tau_1,\tau_2,....\tau_k)\right| = o_p(1).$$

Moreover, it follows from Assumption 6.1(c) and Lemma 2.1 that

$$\begin{aligned}
&\sup_{(\tau_1,\tau_2,....\tau_k)\in\Upsilon^k}\sup_{\theta\in\Theta}\left\|\mathrm{Re}\left[\widehat{b}_{n,k}(\tau_1,\tau_2,....\tau_k|\theta)\right] - E\left(\mathrm{Re}\left[\widehat{b}_{n,k}(\tau_1,\tau_2,....\tau_k|\theta)\right]\right)\right\| \\
&= o_p(1), \\
&\sup_{(\tau_1,\tau_2,....\tau_k)\in\Upsilon^k}\sup_{\theta\in\Theta}\left\|\mathrm{Im}\left[\widehat{b}_{n,k}(\tau_1,\tau_2,....\tau_k|\theta)\right] - E\left(\mathrm{Im}\left[\widehat{b}_{n,k}(\tau_1,\tau_2,....\tau_k|\theta)\right]\right)\right\| \\
&= o_p(1),
\end{aligned}$$

and

$$\sup_{\theta\in\Theta}\left\|\widehat{A}_{2,n}(\theta) - E\left[\widehat{A}_{2,n}(\theta)\right]\right\| = o_p(1),$$

where in the latter case the matrix norm involved is defined by $||A|| = \sqrt{\mathrm{trace}(AA')}$, hence by Lemma 2.1 and Assumption 6.1(h) it follows that

Lemma 6.3. *Under the conditions of Theorem 6.1 and for each k,*

$$\sup_{(\tau_1,\tau_2,....\tau_k)\in\Upsilon^k}\left\|\widehat{A}_{2,n}(\widehat{\theta}_n)^{-1}\widehat{b}_{n,k}(\tau_1,\tau_2,....\tau_k|\widehat{\theta}_n) - A_2^{-1}(\theta_0)b_k(\tau_1,\tau_2,....\tau_k)\right\| = o_p(1).$$

Combining the results of Lemmas 6.1-6.3 it follows that (6.3) holds, which in its turn implies that,

Lemma 6.4. *Under the conditions of Theorem 6.1, $|\widetilde{B}_{i,k,n} - B^*_{i,k,n}| = o_p(1)$ for each $k\in\mathbb{N}$ and each $i = 1,2,...,M$.*

Proof. Denote

$$\delta_{i,k,n} = \sup_{(\tau_1,\tau_2,....\tau_k)\in\Upsilon^k}\left|\widetilde{W}_{i,k,n}(\tau_1,\tau_2,....\tau_k|\widehat{\theta}_n) - W^*_{i,k,n}(\tau_1,\tau_2,....\tau_k)\right|.$$

It is easy to verify that

$$\begin{aligned}
&|\widetilde{B}_{i,k,n} - B^*_{i,k,n}| \\
&\quad \leq \sup_{(\tau_1,\tau_2,....\tau_k)\in\Upsilon^k} \left| \left|\widetilde{W}_{i,k,n}(\tau_1,\tau_2,....\tau_k|\widehat{\theta}_n)\right|^2 - \left|W^*_{i,k,n}(\tau_1,\tau_2,....\tau_k)\right|^2 \right| \\
&\quad \leq \delta^2_{i,k,n} + 2\delta_{i,k,n} \sup_{(\tau_1,\tau_2,....\tau_k)\in\Upsilon^k} \left|W^*_{i,k,n}(\tau_1,\tau_2,....\tau_k)\right|.
\end{aligned}$$

Since by part (a) of Theorem 6.1,

$$W^*_{i,k,n}(\tau_1,\tau_2,....\tau_k) \Rightarrow W^*_{i,k}(\tau_1,\tau_2,....\tau_k)$$

on Υ^k, it follows from the continuous mapping theorem that

$$\sup_{(\tau_1,\tau_2,....\tau_k)\in\Upsilon^k} \left|W^*_{i,k,n}(\tau_1,\tau_2,....\tau_k)\right| \xrightarrow{\mathrm{d}} \sup_{(\tau_1,\tau_2,....\tau_k)\in\Upsilon^k} \left|W^*_{i,k}(\tau_1,\tau_2,....\tau_k)\right|,$$

which implies that

$$\sup_{(\tau_1,\tau_2,....\tau_k)\in\Upsilon^k} \left|W^*_{i,k,n}(\tau_1,\tau_2,....\tau_k)\right| = O_p(1).$$

The lemma now follows straightforwardly from the latter and $\delta_{i,k,n} = o_p(1)$. ■

Lemma 6.5. *Under the conditions of Theorem 6.1,*

$$\sum_{k=1}^{L_n} \gamma_k \widetilde{B}_{i,k,n} - \sum_{k=1}^{\infty} \gamma_k B^*_{i,k,n} = o_p(1),$$

hence

$$(\widetilde{T}_{1,n}, \widetilde{T}_{2,n}, ..., \widetilde{T}_{M,n})' = (T^*_{1,n}, T^*_{2,n}, ..., T^*_{M,n})' + o_p(1),$$

where $T^*_{i,n}$ is defined in Theorem 6.1 and $\widetilde{T}_{i,n}$ in (6.2).

Proof. Note that similar to (3.13), with $\mathcal{F}^{\infty}_{-\infty} = \sigma\left(\{Z_t\}^{\infty}_{t=-\infty}\right)$,

$$\begin{aligned}
&E\left[|\widetilde{W}_{i,k,n}(\tau_1,\tau_2,....\tau_k|\theta)|^2|\mathcal{F}^{\infty}_{-\infty}\right] \\
&= \frac{1}{n}\sum_{t=1}^{n}(Y_t - f_{t-1}(\theta))^2
\end{aligned}$$

$$
\begin{aligned}
&-2\,\mathrm{Re}[\widehat{b}_{n,k}(\tau_1,\tau_2,....\tau_k|\theta)]'\widehat{A}_{2,n}^{-1}(\theta)\,\mathrm{Re}[\widehat{c}_{n,k}(\tau_1,\tau_2,....\tau_k|\theta)]\\
&-2\,\mathrm{Im}[\widehat{b}_{n,k}(\tau_1,\tau_2,....\tau_k|\theta)]'\widehat{A}_{2,n}^{-1}(\theta)\,\mathrm{Im}[\widehat{c}_{n,k}(\tau_1,\tau_2,....\tau_k|\theta)]\\
&+\mathrm{Re}[\widehat{b}_{n,k}(\tau_1,\tau_2,....\tau_k|\theta)]'\widehat{A}_{2,n}^{-1}(\theta)\widehat{A}_{1,n}(\theta)\widehat{A}_{2,n}^{-1}(\theta)\,\mathrm{Re}[\widehat{b}_{n,k}(\tau_1,\tau_2,....\tau_k|\theta)]\\
&+\mathrm{Im}[\widehat{b}_{n,k}(\tau_1,\tau_2,....\tau_k|\theta)]'\widehat{A}_{2,n}^{-1}(\theta)\widehat{A}_{1,n}(\theta)\widehat{A}_{2,n}^{-1}(\theta)\,\mathrm{Im}[\widehat{b}_{n,k}(\tau_1,\tau_2,....\tau_k|\theta)]
\end{aligned}
$$

where

$$
\begin{aligned}
\widehat{c}_{n,k}(\tau_1,\tau_2,....\tau_k|\theta) &= \frac{1}{n}\sum_{t=1}^{n}(Y_t-f_{t-1}(\theta))^2\nabla f_{t-1}(\theta)\exp\left(\mathbf{i}\sum_{j=1}^{k}\tau_j'\Phi(Z_{t-j})\right),\\
\widehat{A}_{1,n}(\theta) &= \frac{1}{n}\sum_{t=1}^{n}(Y_t-f_{t-1}(\theta))^2(\nabla f_{t-1}(\theta))(\nabla f_{t-1}(\theta))'.
\end{aligned}
$$

Then similar to (3.14),

$$
\begin{aligned}
&\int_{\Upsilon^k} E\left[|\widetilde{W}_{i,k,n}(\tau_1,\tau_2,....\tau_k|\theta)|^2|\mathcal{F}_{-\infty}^{\infty}\right]\mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k)\\
&\leq \frac{1}{n}\sum_{t=1}^{n}(Y_t-f_{t-1}(\theta))^2+4.\mathrm{trace}(\widehat{A}_{2,n}^{-1}(\theta))\sqrt{\mathrm{trace}(\widehat{A}_{2,n}(\theta))}\\
&+4.\mathrm{trace}(\widehat{A}_{2,n}^{-1}(\theta))\sqrt{\mathrm{trace}(\widehat{A}_{2,n}(\theta))}\frac{1}{n}\sum_{t=1}^{n}(Y_t-f_{t-1}(\theta))^2||\nabla f_{t-1}(\theta)||\\
&+2.\mathrm{trace}\left(\widehat{A}_{2,n}^{-1}(\theta)\widehat{A}_{1,n}(\theta)\widehat{A}_{2,n}^{-1}(\theta)\right)\mathrm{trace}\left(\widehat{A}_{2,n}(\theta)\right)\\
&=\widehat{R}_n(\theta),
\end{aligned}
$$

say, hence $E\left[\widetilde{B}_{i,k,n}|\mathcal{F}_{-\infty}^{\infty}\right]\leq \widehat{R}_n(\widehat{\theta}_n)$ and thus for $K\in\mathbb{N}$,

$$
E\left[\sum_{k=K}^{\infty}\gamma_k\widetilde{B}_{i,k,n}\middle|\mathcal{F}_{-\infty}^{\infty}\right]\leq \widehat{R}_n(\widehat{\theta}_n)\sum_{k=K}^{\infty}\gamma_k.
$$

It is not hard (but somewhat tedious) to verify from Assumption 6.1 and Lemma 2.1 that $\sup_{\theta\in\Theta}\left|\widehat{R}_n(\theta)-R(\theta)\right|=o_p(1)$ and thus by Lemma 2.1,

$$
\widehat{R}_n(\widehat{\theta}_n)\xrightarrow{\mathrm{p}} R(\theta_0),
$$

where

$$
\begin{aligned}
R(\theta) &= E[(Y_t - f_{t-1}(\theta))^2] + 4.\text{trace}(A_2^{-1}(\theta))\sqrt{\text{trace}(A_2(\theta))} \\
&\quad +4.\text{trace}(A_2^{-1}(\theta))\sqrt{\text{trace}(A_2(\theta))}E\left[(Y_t - f_{t-1}(\theta))^2 ||\nabla f_{t-1}(\theta)||\right] \\
&\quad +2.\text{trace}\left(A_2^{-1}(\theta)A_1(\theta)A_2^{-1}(\theta)\right)\text{trace}\left(A_2(\theta)\right)
\end{aligned}
$$

with $A_2(\theta)$ defined in Assumption 6.1(h) and

$$
A_1(\theta) = E\left[(Y_t - f_{t-1}(\theta))^2(\nabla f_{t-1}(\theta))(\nabla f_{t-1}(\theta))'\right].
$$

Now by Chebyshev's inequality for conditional probabilities and first moments, it follows that for arbitrary $\beta > 0$,

$$
\Pr\left[\sum_{k=K}^{\infty}\gamma_k\widetilde{B}_{i,k,n} > \beta.\widehat{R}_n(\widehat{\theta}_n)\middle| \mathcal{F}_{-\infty}^{\infty}\right] \leq \frac{E\left[\sum_{k=K}^{\infty}\gamma_k\widetilde{B}_{i,k,n}|\mathcal{F}_{-\infty}^{\infty}\right]}{\beta.\widehat{R}_n(\widehat{\theta}_n)} \leq \frac{\sum_{k=K}^{\infty}\gamma_k}{\beta},
$$

hence

$$
\begin{aligned}
&\frac{\sum_{k=K}^{\infty}\gamma_k}{\beta} \\
&\geq E\left(\Pr\left[\sum_{k=K}^{\infty}\gamma_k\widetilde{B}_{i,k,n} > \beta.\widehat{R}_n(\widehat{\theta}_n)\middle| \mathcal{F}_{-\infty}^{\infty}\right]\right) \\
&= \Pr\left[\sum_{k=K}^{\infty}\gamma_k\widetilde{B}_{i,k,n} > \beta.\widehat{R}_n(\widehat{\theta}_n)\right] \\
&= \Pr\left[\sum_{k=K}^{\infty}\gamma_k\widetilde{B}_{i,k,n} > \beta.\widehat{R}_n(\widehat{\theta}_n) \text{ and } |\widehat{R}_n(\widehat{\theta}_n) - R(\theta_0)| > 1\right] \\
&+\Pr\left[\sum_{k=K}^{\infty}\gamma_k\widetilde{B}_{i,k,n} > \beta.\widehat{R}_n(\widehat{\theta}_n) \text{ and } |\widehat{R}_n(\widehat{\theta}_n) - R(\theta_0)| \leq 1\right] \\
&\geq \Pr\left[\sum_{k=K}^{\infty}\gamma_k\widetilde{B}_{i,k,n} > \beta.\widehat{R}_n(\widehat{\theta}_n) \text{ and } |\widehat{R}_n(\widehat{\theta}_n) - R(\theta_0)| > 1\right] \\
&+\Pr\left[\sum_{k=K}^{\infty}\gamma_k\widetilde{B}_{i,k,n} > \beta.(R(\theta_0)+1) \text{ and } |\widehat{R}_n(\widehat{\theta}_n) - R(\theta_0)| \leq 1\right] \\
&= \Pr\left[\sum_{k=K}^{\infty}\gamma_k\widetilde{B}_{i,k,n} > \beta.\widehat{R}_n(\widehat{\theta}_n) \text{ and } |\widehat{R}_n(\widehat{\theta}_n) - R(\theta_0)| > 1\right]
\end{aligned}
$$

$$+\Pr\left[\sum_{k=K}^{\infty}\gamma_k\widetilde{B}_{i,k,n} > \beta.(R(\theta_0)+1)\right]$$
$$-\Pr\left[\sum_{k=K}^{\infty}\gamma_k\widetilde{B}_{i,k,n} > \beta.(R(\theta_0)+1) \text{ and } |\widehat{R}_n(\widehat{\theta}_n)-R(\theta_0)|>1\right] \quad (6.9)$$

Since

$$\limsup_{n\to\infty}\Pr\left[\sum_{k=K}^{\infty}\gamma_k\widetilde{B}_{i,k,n} > \beta.\widehat{R}_n(\widehat{\theta}_n) \text{ and } |\widehat{R}_n(\widehat{\theta}_n)-R(\theta_0)|>1\right]$$
$$\leq \limsup_{n\to\infty}\Pr\left[|\widehat{R}_n(\widehat{\theta}_n)-R(\theta_0)|>1\right]=0$$

and

$$\limsup_{n\to\infty}\Pr\left[\sum_{k=K}^{\infty}\gamma_k\widetilde{B}_{i,k,n} > \beta.(R(\theta_0)+1) \text{ and } |\widehat{R}_n(\widehat{\theta}_n)-R(\theta_0)|>1\right]$$
$$\leq \limsup_{n\to\infty}\Pr\left[|\widehat{R}_n(\widehat{\theta}_n)-R(\theta_0)|>1\right]=0$$

it follows now from (6.9) that for arbitrary $\beta>0$,

$$\limsup_{n\to\infty}\Pr\left[\sum_{k=K}^{\infty}\gamma_k\widetilde{B}_{i,k,n} > \beta.(R(\theta_0)+1)\right] \leq \frac{\sum_{k=K}^{\infty}\gamma_k}{\beta}$$

For arbitrary $\varepsilon>0$ and $\delta\in(0,1)$, choose $\beta=\varepsilon/(R(\theta_0)+1)$ and let $K_0(\varepsilon,\delta)\in\mathbb{N}$ be such that $\sum_{k=K_0(\varepsilon,\delta)}^{\infty}\gamma_k<\delta.\beta$. Then for $K\geq K_0(\varepsilon,\delta)$,

$$\limsup_{n\to\infty}\Pr\left[\sum_{k=K}^{\infty}\gamma_k\widetilde{B}_{i,k,n} > \varepsilon\right] < \delta. \quad (6.10)$$

Note that similar to (3.14), $E[B^*_{i,k,n}]\leq R(\theta_0)$, so that $K_0(\varepsilon,\delta)$ can be chosen such that for $K\geq K_0(\varepsilon,\delta)$,

$$\limsup_{n\to\infty}\Pr\left[\sum_{k=K}^{\infty}\gamma_k B^*_{i,k,n} > \varepsilon\right] < \delta. \quad (6.11)$$

It follows now from (6.10), (6.11) and Lemma 6.4, similar to the proof of Lemma 3.4, that $\sum_{k=1}^{\infty}\gamma_k\widetilde{B}_{i,k,n}-\sum_{k=1}^{\infty}\gamma_k B^*_{i,k,n}=o_p(1)$ and $\sum_{k=L_n+1}^{\infty}\gamma_k\widetilde{B}_{i,k,n}=o_p(1)$. ■

Summarizing, the following results have been proved.

Theorem 6.2. *Under the conditions and notations in Theorem 6.1 it follows that*

$$(\widetilde{T}_{1,n}, \widetilde{T}_{2,n}, ..., \widetilde{T}_{M,n})' \xrightarrow{\mathrm{d}} (T_1^*, T_2^*, ..., T_M^*)',$$

where $T_1^*, T_2^*, ..., T_M^*$ *are i.i.d.. Moreover, with* $\widehat{T}_n$ *the WICM test statistic, it follows that under* H_0,

$$(\widehat{T}_n, \widetilde{T}_{1,n}, \widetilde{T}_{2,n}, ..., \widetilde{T}_{M,n})' \xrightarrow{\mathrm{d}} (T, T_1, T_2, ..., T_M)',$$

where $T, T_1, T_2, ..., T_M$ *are i.i.d.*

Again, the asymptotic $\alpha \times 100\%$ critical values $c_0(\alpha)$ of the WICM test, i.e., $\Pr[T > c_0(\alpha)] = \alpha$, can be approximated by sorting the $\widetilde{T}_{i,n}$'s in decreasing order and then use $\widetilde{c}_{n,M}(\alpha) = \widetilde{T}_{[\alpha M],n}$ as the approximation of $c_0(\alpha)$. More precisely,

Theorem 6.3. *Let* $\widetilde{c}_{n,M}(\alpha)$ *be the* $1-\alpha$ *quantile of the empirical distribution function*

$$\widetilde{G}_{n,M}(x) = \frac{1}{M}\sum_{i=1}^{M} I\left(\widetilde{T}_{i,n} \le x\right),$$

i.e., $\widetilde{c}_{n,M}(\alpha) = \arg\min_{\widetilde{G}_{n,M}(x) \ge 1-\alpha} x$, *and let* $c(\alpha)$ *be the* $1-\alpha$ *quantile of the distribution functions* $G(x) = \Pr[T \le x]$. *Then under* H_0, *the conditions of Theorem 6.2 and for arbitrary* $\delta > 0$,

$$\limsup_{n\to\infty} \Pr\left[|\widetilde{G}_{n,M}(x) - G(x)| > \delta\right] \le \frac{1}{4\delta M}$$

hence

$$\lim_{M\to\infty} \limsup_{n\to\infty} \Pr[|\widetilde{c}_{n,M}(\alpha) - c(\alpha)| > \delta] = 0.$$

Moreover, under H_1, $\Pr[\widehat{T}_n > \widetilde{c}_{n,M}(\alpha)] \to 1$ *for any* $M \in \mathbb{N}$.

Proof. Similar to Theorems 6.2 and 6.3 in the addendum to Bierens (1982) in Chapter 2, using the fact that by Theorem 5.3, it follows that $G(x)$ is continuous on $(0, \infty)$, so that for all $\alpha \in (0, 1)$, $G(c(\alpha)) = 1 - \alpha$.

Under H_1 we have for arbitrary $\delta > 0$

$$\limsup_{n\to\infty} \Pr\left[|\widetilde{G}_{n,M}(x) - G^*(x)| > \delta\right] \leq \frac{1}{4\delta M}$$

where $G^*(x) = \Pr[T_1^* \leq x]$, and

$$\lim_{M\to\infty} \limsup_{n\to\infty} \Pr[|\widetilde{c}_{n,M}(\alpha) - c^*(\alpha)| > \delta] = 0$$

where $c^*(\alpha)$ is the $1-\alpha$ quantile of $G^*(x)$, but these results are of no direct use, except that they imply that under H_1, $\widetilde{c}_{n,M}(\alpha) = O_p(1)$ and thus $\Pr[\widehat{T}_n > \widetilde{c}_{n,M}(\alpha)] \to 1$, for any $M \in \mathbb{N}$. ■

7. The initial value problem

The results in this addendum were based on the assumption that all lagged Z_t's are observed, which is obviously not realistic. In particular, given that Z_t is observed for $1 - t_0 \leq t \leq n$, where $t_0 \geq \max(p, q)$, the empirical counter-part of $\psi_k(\tau_1, \tau_2, ...\tau_k)$ defined in (1.10) is

$$\begin{aligned}\widetilde{\psi}_{k,n}(\tau_1, \tau_2, ...\tau_k) &= \frac{1}{n}\sum_{t=1}^{n}(Y_t - f_{t-1}(\widehat{\theta}_n)) \exp\left(\mathbf{i}.\sum_{j=1}^{\min(k,t-t_0)} \tau_j'\Phi(Z_{t-j})\right) \\ &= \widetilde{W}_{k,n}(\tau_1, \tau_2, ...\tau_k)/\sqrt{n}\end{aligned}$$

whereas all the asymptotic results were based on

$$\begin{aligned}\widehat{\psi}_{k,n}(\tau_1, \tau_2, ...\tau_k) &= \frac{1}{n}\sum_{t=1}^{n}(Y_t - f_{t-1}(\widehat{\theta}_n)) \exp\left(\mathbf{i}.\sum_{j=1}^{k} \tau_j'\Phi(Z_{t-j})\right) \\ &= \widehat{W}_{k,n}(\tau_1, \tau_2, ...\tau_k)/\sqrt{n}\end{aligned}$$

As alluded to in the introduction, we need to set forth conditions under which the asymptotic results above on the basis of $\widehat{W}_{k,n}$ are the same as for $\widetilde{W}_{k,n}$. The condition involved is condition (7.1) in the following theorem.

Theorem 7.1. *Denote*

$$\widetilde{B}_{n,k} = \int_{\Upsilon^k} \left|\widetilde{W}_{k,n}(\tau_1, \tau_2, ...\tau_k)\right|^2 \mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k),$$

$$\widehat{B}_{n,k} = \int_{\Upsilon^k} \left|\widehat{W}_{k,n}(\tau_1, \tau_2, ...\tau_k)\right|^2 \mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k).$$

If the sequence of weights γ_k is chosen such that

$$\sum_{k=1}^{\infty} k^2 \sqrt{\gamma_k} < \infty \tag{7.1}$$

then under H_0 and the conditions and notations in Theorem 3.1,

$$\sum_{k=1}^{L_n} \gamma_k \widetilde{B}_{n,k} = \sum_{k=1}^{L_n} \gamma_k \widehat{B}_{n,k} + O_p(1/\sqrt{n}),$$

whereas under H_1,

$$\sum_{k=1}^{L_n} \gamma_k \widetilde{B}_{n,k}/n = \sum_{k=1}^{L_n} \gamma_k \widehat{B}_{n,k}/n + O_p(1/n),$$

Proof. Observe that

$$\begin{aligned}
&\left| \exp\left(\mathbf{i}. \sum_{j=1}^{k} \tau_j' \Phi(Z_{t-j}) \right) - \exp\left(\mathbf{i}. \sum_{j=1}^{\min(k,t-t_0)} \tau_j' \Phi(Z_{t-j}) \right) \right| \\
&\qquad \leq \left| \sum_{j=\min(k,t-t_0)+1}^{k} \tau_j' \Phi(Z_{t-j}) \right| \\
&\qquad = 0 \text{ for } t \geq k + t_0 \\
&\qquad \leq C(k + t_0 - 1) \text{ for } 1 \leq t < k + t_0
\end{aligned}$$

where $C = \sup_{\tau \in \Upsilon} ||\tau||. \sup_{z \in \mathbb{R}^{s+1}} ||\Phi(z)||$. Hence,

$$\left| \widetilde{W}_{k,n}(\tau_1, \tau_2, ...\tau_k) - \widehat{W}_{k,n}(\tau_1, \tau_2, ...\tau_k) \right| \leq C. \frac{1}{\sqrt{n}} \sum_{t=1}^{k+t_0-1} |Y_t - f_{t-1}(\widehat{\theta}_n)|(k + t_0 - 1)$$

and thus

$$\begin{aligned}
&\left| |\widetilde{W}_{k,n}(\tau_1, \tau_2, ...\tau_k)|^2 - |\widehat{W}_{k,n}(\tau_1, \tau_2, ...\tau_k)|^2 \right| \\
&\qquad \leq \left| \widetilde{W}_{k,n}(\tau_1, \tau_2, ...\tau_k) - \widehat{W}_{k,n}(\tau_1, \tau_2, ...\tau_k) \right|^2 \\
&\qquad + 2\sqrt{2} \left| \widetilde{W}_{k,n}(\tau_1, \tau_2, ...\tau_k) - \widehat{W}_{k,n}(\tau_1, \tau_2, ...\tau_k) \right| \times |\widehat{W}_{k,n}(\tau_1, \tau_2, ...\tau_k) |
\end{aligned}$$

$$\leq C^2\left(\frac{1}{\sqrt{n}}\sum_{t=1}^{k+t_0-1}|Y_t - f_{t-1}(\widehat{\theta}_n)|\right)^2 (k+t_0-1)^2$$

$$+\, 2\sqrt{2}C\frac{1}{\sqrt{n}}\sum_{t=1}^{k+t_0-1}|Y_t - f_{t-1}(\widehat{\theta}_n)|(k+t_0-1)\times|\widehat{W}_{k,n}(\tau_1,\tau_2,....\tau_k)|,$$

where the first inequality follow from the easy inequality $||z_1|^2 - |z_2|^2| \leq |z_1 - z_2|^2 + 2\sqrt{2}|z_1-z_2|.|z_2|$ for complex numbers z_1 and z_2. Consequently,

$$\begin{aligned}
\left|\widetilde{B}_{n,k} - \widehat{B}_{n,k}\right| &\leq \frac{1}{n}C^2(k+t_0-1)^4\left(\frac{1}{k+t_0-1}\sum_{t=1}^{k+t_0-1}|Y_t - f_{t-1}(\widehat{\theta}_n)|\right)^2 \\
&\quad + \frac{1}{\sqrt{n}}2\sqrt{2}C(k+t_0-1)\sum_{t=1}^{k+t_0-1}|Y_t - f_{t-1}(\widehat{\theta}_n)|\times\sqrt{\widehat{B}_{n,k}} \\
&\leq \frac{1}{n}C^2(k+t_0-1)^4\frac{1}{k+t_0-1}\sum_{t=1}^{k+t_0-1}(Y_t - f_{t-1}(\widehat{\theta}_n))^2 \\
&\quad + \frac{1}{\sqrt{n}}2\sqrt{2}C(k+t_0-1)\sum_{t=1}^{k+t_0-1}|Y_t - f_{t-1}(\widehat{\theta}_n)|\times\sqrt{\widehat{B}_{n,k}}
\end{aligned}$$

which implies that

$$\begin{aligned}
&\left|\sum_{k=1}^{L_n}\gamma_k\widetilde{B}_{n,k} - \sum_{k=1}^{L_n}\gamma_k\widehat{B}_{n,k}\right| \\
&\leq \frac{1}{n}C^2\sum_{k=1}^{L_n}\gamma_k(k+t_0-1)^4\left(\frac{1}{k+t_0-1}\sum_{t=1}^{k+t_0-1}\sup_{\theta\in\Theta}(Y_t - f_{t-1}(\theta))^2\right) \\
&+\frac{1}{\sqrt{n}}2\sqrt{2}C\sum_{k=1}^{L_n}\sqrt{\gamma_k}(k+t_0-1)^2\left(\frac{1}{k+t_0-1}\sum_{t=1}^{k+t_0-1}\sup_{\theta\in\Theta}|Y_t - f_{t-1}(\theta)|\right) \\
&\times\sqrt{\sum_{k=1}^{L_n}\gamma_k\widehat{B}_{n,k}}
\end{aligned}$$

Next, note that (7.1) implies

$$\sum_{k=1}^{\infty}k^4\gamma_k < \infty \tag{7.2}$$

Since by Assumption 1.1 and (7.1) and (7.2),

$$E\left[\sum_{k=1}^{L_n}\gamma_k(k+t_0-1)^4\left(\frac{1}{k+t_0-1}\sum_{t=1}^{k+t_0-1}\sup_{\theta\in\Theta}(Y_t-f_{t-1}(\theta))^2\right)\right]$$
$$=E\left[\sup_{\theta\in\Theta}(Y_t-f_{t-1}(\theta))^2\right]\sum_{k=1}^{\infty}\gamma_k(k+t_0-1)^4<\infty$$

and

$$E\left[\sum_{k=1}^{L_n}\sqrt{\gamma_k}(k+t_0-1)^2\left(\frac{1}{k+t_0-1}\sum_{t=1}^{k+t_0-1}\sup_{\theta\in\Theta}|Y_t-f_{t-1}(\theta)|\right)\right]$$
$$=E\left[\sup_{\theta\in\Theta}|Y_t-f_{t-1}(\theta)|\right]\sum_{k=1}^{\infty}\sqrt{\gamma_k}(k+t_0-1)^2<\infty,$$

whereas by Theorem 3.1, $\sum_{k=1}^{L_n}\gamma_k\widehat{B}_{n,k}\xrightarrow{\mathrm{d}}\sum_{k=1}^{\infty}\gamma_k B_k$ under H_0, which implies that $\sum_{k=1}^{L_n}\gamma_k\widehat{B}_{n,k}=O_p(1)$, it follows now that under H_0,

$$\left|\sum_{k=1}^{L_n}\gamma_k\widetilde{B}_{n,k}-\sum_{k=1}^{L_n}\gamma_k\widehat{B}_{n,k}\right|=O_p(1/\sqrt{n}).$$

Finally, the result under H_1 follows easily from

$$\left|\sum_{k=1}^{L_n}\gamma_k\widetilde{B}_{n,k}/n-\sum_{k=1}^{L_n}\gamma_k\widehat{B}_{n,k}/n\right|$$
$$\leq\frac{1}{n^2}C^2\sum_{k=1}^{L_n}\gamma_k(k+t_0-1)^4\left(\frac{1}{k+t_0-1}\sum_{t=1}^{k+t_0-1}\sup_{\theta\in\Theta}(Y_t-f_{t-1}(\theta))^2\right)$$
$$+\frac{1}{n}2\sqrt{2}C\sum_{k=1}^{L_n}\sqrt{\gamma_k}(k+t_0-1)^2\left(\frac{1}{k+t_0-1}\sum_{t=1}^{k+t_0-1}\sup_{\theta\in\Theta}|Y_t-f_{t-1}(\theta)|\right)$$
$$\times\sqrt{\sum_{k=1}^{L_n}\gamma_k\widehat{B}_{n,k}/n}$$

and Theorem 4.1. ■

8. Standardization of the conditioning variables

Another unresolved issue in B84 is how to standardize the conditioning lagged variables in Z_t before taking the bounded transformation Φ in order to preserve enough variation in $\Phi(Z_t)$. See section 6 in B84.

To discuss this issue, let us for the time being assume that $Z_t = Y_t$, so that the nonlinear ARX model (1.1) becomes a nonlinear AR model:

$$Y_t = f(Y_{t-1}, Y_{t-2}, ..., Y_{t-p}, \theta_0) + U_t = f_{t-1}(\theta_0) + U_t, \text{ say.}$$

Moreover, as in B82, let us assume that

$$\Phi(x) = \arctan(x), \tag{8.1}$$

and note that

$$\Phi'(x) = \left(1 + x^2\right)^{-1}. \tag{8.2}$$

8.1. A wrong way to standardize

Following B82, one may be tempted to standardize each lagged Y_t as $\widetilde{Y}_{n,t} = \widehat{\sigma}_n^{-1}(Y_t - \widehat{\mu}_n) = \widehat{\alpha}_n Y_t - \widehat{\beta}_n$, where

$$\begin{aligned} \widehat{\mu}_n &= (1/(n+t_0)) \sum_{t=1-t_0}^{n} Y_t, \ \widehat{\sigma}_n = \sqrt{(1/(n+t_0-1)) \sum_{t=1-t_0}^{n} (Y_t - \widehat{\mu}_n)^2}, \\ \widehat{\alpha}_n &= 1/\widehat{\sigma}_n, \ \widehat{\beta}_n = \widehat{\mu}_n / \widehat{\sigma}_n, \end{aligned}$$

and $1 - t_0$ is the time index of the first observed Y_t. Then (1.13) becomes

$$\widetilde{W}_{k,n}(\tau_1, \tau_2, ...\tau_k) = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} (Y_t - f_{t-1}(\widehat{\theta}_n)) \exp\left(\mathbf{i}. \sum_{j=1}^{k} \tau_j \Phi\left(\widehat{\alpha}_n Y_{t-j} - \widehat{\beta}_n\right)\right). \tag{8.3}$$

Similarly, denoting

$$\mu_0 = E[Y_t], \ \sigma_0 = \sqrt{E[(Y_t - \mu_0)^2]}, \ \alpha_0 = 1/\sigma_0, \ \beta_0 = \mu_0/\sigma_0, \tag{8.4}$$

replace $\Phi(Y_{t-j})$ in (1.13) by $\Phi\left(\alpha_0 Y_{t-j} - \beta_0\right)$, i.e.,

$$\widehat{W}_{k,n}(\tau_1, \tau_2, ...\tau_k) = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} (Y_t - f_{t-1}(\widehat{\theta}_n)) \exp\left(\mathbf{i}. \sum_{j=1}^{k} \tau_j \Phi\left(\alpha_0 Y_{t-j} - \beta_0\right)\right). \tag{8.5}$$

Now the hope is that under H_0, $\widetilde{W}_{k,n}(\tau_1, \tau_2, ...\tau_k)$ and $\widehat{W}_{k,n}(\tau_1, \tau_2, ...\tau_k)$ are asymptotically equivalent, in the sense the WICM test based $\widetilde{W}_{k,n}(\tau_1, \tau_2, ...\tau_k)$ has the same limiting null distribution as the (now infeasible) WICM test based on $\widehat{W}_{k,n}(\tau_1, \tau_2, ...\tau_k)$.

It can be shown that it does if $\widehat{\alpha}_n = \alpha_0 + O_p(n^{-1/2})$ and $\widehat{\beta}_n = \beta_0 + O_p(n^{-1/2})$, but it is unlikely that these conditions hold in the time series case. Moreover, even if it does (which is doubtful), this particular standardization is not recommended because it destroys the martingale difference structure of the first term in (8.3) under the null hypothesis and therefore may lead to rejection of the true null hypothesis in the case of small and medium size macro-economics data.

8.2. Martingale difference structure preserving standardization

Therefore, a better standardization procedure is the following. Denote

$$\begin{aligned}
\widehat{\mu}_t &= (t+t_0)^{-1} \sum_{i=1-t_0}^{t} Y_i,\ \widehat{\sigma}_t = \sqrt{(t+t_0)^{-1} \sum_{i=1-t_0}^{t} (Y_i - \widehat{\mu}_t)^2},\ \text{if } t > 1 - t_0, \\
\widehat{\mu}_t &= 0,\ \widehat{\sigma}_t = 1 \text{ if } t \leq 1 - t_0, \\
\widehat{\alpha}_t &= 1/\widehat{\sigma}_t,\ \widehat{\beta}_t = \widehat{\mu}_t/\widehat{\sigma}_t, \\
\underline{\widetilde{Y}}_t &= \widehat{\alpha}_t Y_t - \widehat{\beta}_t,\ \underline{Y}_t = \alpha_0 Y_t - \beta_0,
\end{aligned} \tag{8.6}$$

where α_0 and β_0 are the same as in (8.4). Then

Lemma 8.1 *Under Assumptions 1.1(a) and 2.1(a) and with* Φ *as in* (8.1), *it follows that*

$$p \lim_{t\to\infty} \left| \Phi\left(\underline{\widetilde{Y}}_t\right) - \Phi\left(\underline{Y}_t\right) \right| = 0. \tag{8.7}$$

Consequently, for an arbitrary strictly stationary time series process $V_t \in \mathbb{R}$ *satisfying* $E[|V_t|] < \infty$ *and each* $j \in \mathbb{N}$,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{t=1}^{n} E\left[|V_t|. \left| \Phi\left(\underline{\widetilde{Y}}_{t-j}\right) - \Phi\left(\underline{Y}_{t-j}\right) \right| \right] = 0, \tag{8.8}$$

which by Chebyshev's inequality for first moments implies that

$$p \lim_{n\to\infty} \frac{1}{n} \sum_{t=1}^{n} |V_t|. \left| \Phi\left(\underline{\widetilde{Y}}_{t-j}\right) - \Phi\left(\underline{Y}_{t-j}\right) \right| = 0.$$

Proof. Recall from Assumptions 1.1(a) and 2.1(a) that Y_t and Y_t^2 are strictly stationary with vanishing memory, and $E[Y_t^2] < \infty$, so that by Lemma 2.2, $p\lim_{t\to\infty} \widehat{\mu}_t = E[Y_1] = \mu_0$, $p\lim_{t\to\infty} \widehat{\sigma}_t^2 = \mathrm{var}(Y_1) = \sigma_0^2$, hence

$$p \lim_{t\to\infty} \widehat{\alpha}_t = \alpha_0, \; p \lim_{t\to\infty} \widehat{\beta}_t = \beta_0. \tag{8.9}$$

Moreover, with Φ as in (8.1), it follows from the mean value theorem and (8.2) that

$$\left|\Phi\left(\underline{\widetilde{Y}}_t\right) - \Phi\left(\underline{Y}_t\right)\right| \le \left|\underline{\widetilde{Y}}_t - \underline{Y}_t\right| \le |\widehat{\alpha}_t - \alpha_0|.|Y_t| + |\widehat{\beta}_t - \beta_0|$$

so that (8.7) holds by (8.9) and the fact that by strict stationarity, $|Y_t| = O_p(1)$.

As to (8.8), observe that for an arbitrary $M > 0$, and with $I(.)$ the indicator function,

$$\begin{aligned}
&\frac{1}{n}\sum_{t=1}^{n} E\left[|V_t|.\left|\Phi\left(\underline{\widetilde{Y}}_t\right) - \Phi\left(\underline{Y}_t\right)\right|\right] \\
&\quad\le \frac{1}{n}\sum_{t=1}^{n} E\left[|V_t|.I\left(|V_t| \le M\right)\left|\Phi\left(\underline{\widetilde{Y}}_t\right) - \Phi\left(\underline{Y}_t\right)\right|\right] \\
&\qquad + \frac{1}{n}\sum_{t=1}^{n} E\left[|V_t|.I\left(|V_t| > M\right)\left|\Phi\left(\underline{\widetilde{Y}}_t\right) - \Phi\left(\underline{Y}_t\right)\right|\right] \\
&\quad\le M.\frac{1}{n}\sum_{t=1}^{n} E\left[\left|\Phi\left(\underline{\widetilde{Y}}_t\right) - \Phi\left(\underline{Y}_t\right)\right|\right] + \pi.E\left[|V_1|.I\left(|V_1| > M\right)\right],
\end{aligned}$$

where the π follows from the fact that by (8.1), $\sup_{x\in\mathbb{R}} |\Phi(x)| = \pi/2$. For an arbitrary $\varepsilon > 0$ we can choose M so large that $\pi.E\left[|V_1|.I\left(|V_1| > M\right)\right] < \varepsilon$. Moreover, it follows from the bounded convergence theorem and (8.7) that

$$\lim_{t\to\infty} E\left[\left|\Phi\left(\underline{\widetilde{Y}}_t\right) - \Phi\left(\underline{Y}_t\right)\right|\right] = 0,$$

which trivially implies that $\lim_{n\to\infty} \frac{1}{n}\sum_{t=1}^{n} E\left[\left|\Phi\left(\underline{\widetilde{Y}}_t\right) - \Phi\left(\underline{Y}_t\right)\right|\right] = 0$. Hence,

$$\limsup_{n\to\infty} E\left[\frac{1}{n}\sum_{t=1}^{n} |V_t|.\left|\Phi\left(\underline{\widetilde{Y}}_t\right) - \Phi\left(\underline{Y}_t\right)\right|\right] < \varepsilon,$$

which, by the arbitrariness of $\varepsilon > 0$ implies that (8.8) holds. ■

Now (8.5) reads

$$\begin{aligned}\widehat{W}_{k,n}(\tau_1,\tau_2,....\tau_k) &= \frac{1}{\sqrt{n}}\sum_{t=1}^{n}(Y_t - f_{t-1}(\widehat{\theta}_n))\exp\left(\mathbf{i}.\sum_{j=1}^{k}\tau_j\Phi\left(\underline{Y}_{t-j}\right)\right)\\ &= \widehat{W}^{(1)}_{k,n}(\tau_1,\tau_2,....\tau_k) - \widehat{W}^{(2)}_{k,n}(\tau_1,\tau_2,....\tau_k)\end{aligned}$$

where

$$\widehat{W}^{(1)}_{k,n}(\tau_1,\tau_2,....\tau_k) = \frac{1}{\sqrt{n}}\sum_{t=1}^{n}U_t\exp\left(\mathbf{i}.\sum_{j=1}^{k}\tau_j\Phi\left(\underline{Y}_{t-j}\right)\right), \tag{8.10}$$

$$\begin{aligned}&\widehat{W}^{(2)}_{k,n}(\tau_1,\tau_2,....\tau_k)\\ &\quad = \frac{1}{\sqrt{n}}\sum_{t=1}^{n}\left(f_{t-1}(\widehat{\theta}_n) - f_{t-1}(\theta_0)\right)\exp\left(\mathbf{i}.\sum_{j=1}^{k}\tau_j\Phi\left(\underline{Y}_{t-j}\right)\right).\end{aligned} \tag{8.11}$$

Similarly, (8.3) now becomes

$$\begin{aligned}\widetilde{W}_{k,n}(\tau_1,\tau_2,....\tau_k) &= \frac{1}{\sqrt{n}}\sum_{t=1}^{n}(Y_t - f_{t-1}(\widehat{\theta}_n))\exp\left(\mathbf{i}.\sum_{j=1}^{k}\tau_j\Phi\left(\underline{\widetilde{Y}}_{t-j}\right)\right)\\ &= \widetilde{W}^{(1)}_{k,n}(\tau_1,\tau_2,....\tau_k) - \widetilde{W}^{(2)}_{k,n}(\tau_1,\tau_2,....\tau_k)\end{aligned}$$

where

$$\widetilde{W}^{(1)}_{k,n}(\tau_1,\tau_2,....\tau_k) = \frac{1}{\sqrt{n}}\sum_{t=1}^{n}U_t\exp\left(\mathbf{i}.\sum_{j=1}^{k}\tau_j\Phi\left(\underline{\widetilde{Y}}_{t-j}\right)\right), \tag{8.12}$$

$$\begin{aligned}&\widetilde{W}^{(2)}_{k,n}(\tau_1,\tau_2,....\tau_k)\\ &\quad = \frac{1}{\sqrt{n}}\sum_{t=1}^{n}\left(f_{t-1}(\widehat{\theta}_n) - f_{t-1}(\theta_0)\right)\exp\left(\mathbf{i}.\sum_{j=1}^{k}\tau_j\Phi\left(\underline{\widetilde{Y}}_{t-j}\right)\right).\end{aligned} \tag{8.13}$$

8.2.1. The case H_0

Using Lemma 8.1, I will show now that

Lemma 8.2. *Under the conditions of Theorem 2.2, and with* Φ *the* arctan *function,*

$$\sup_{(\tau_1,\tau_2,....\tau_k)'\in\Upsilon^k}\left|\widetilde{W}^{(2)}_{k,n}(\tau_1,\tau_2,....\tau_k)-\widehat{W}^{(2)}_{k,n}(\tau_1,\tau_2,....\tau_k)\right|=o_p(1).$$

Proof. It follows from the mean value theorem that $f_{t-1}(\widehat{\theta}_n)-f_{t-1}(\theta_0)=\left(\widehat{\theta}_n-\theta_0\right)'\nabla f_{t-1}(\widetilde{\theta}_t)$, where $\widetilde{\theta}_t$ is a mean value satisfying $||\widetilde{\theta}_t-\theta_0||\le||\widehat{\theta}_n-\theta_0||$, hence

$$\begin{aligned}
&\left|\mathrm{Re}\left[\widetilde{W}^{(2)}_{k,n}(\tau_1,\tau_2,....\tau_k)\right]-\mathrm{Re}\left[\widehat{W}^{(2)}_{k,n}(\tau_1,\tau_2,....\tau_k)\right]\right|\\
&=\left|\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\left(f_{t-1}(\widehat{\theta}_n)-f_{t-1}(\theta_0)\right)\right.\\
&\qquad\left.\times\left(\cos\left(\sum_{j=1}^{k}\tau_j\Phi\left(\underline{\widetilde{Y}}_{t-j}\right)\right)-\cos\left(\sum_{j=1}^{k}\tau_j\Phi\left(\underline{Y}_{t-j}\right)\right)\right)\right|\\
&\le\left|\sqrt{n}\left(\widehat{\theta}_n-\theta_0\right)'\frac{1}{n}\sum_{t=1}^{n}\left(\nabla f_{t-1}(\widetilde{\theta}_t)-\nabla f_{t-1}(\theta_0)\right)\right.\\
&\qquad\left.\times\left(\cos\left(\sum_{j=1}^{k}\tau_j\Phi\left(\underline{\widetilde{Y}}_{t-j}\right)\right)-\cos\left(\sum_{j=1}^{k}\tau_j\Phi\left(\underline{Y}_{t-j}\right)\right)\right)\right|\\
&+\left|\sqrt{n}\left(\widehat{\theta}_n-\theta_0\right)'\frac{1}{n}\sum_{t=1}^{n}\nabla f_{t-1}(\theta_0)\right.\\
&\qquad\left.\times\left(\cos\left(\sum_{j=1}^{k}\tau_j\Phi\left(\underline{\widetilde{Y}}_{t-j}\right)\right)-\cos\left(\sum_{j=1}^{k}\tau_j\Phi\left(\underline{Y}_{t-j}\right)\right)\right)\right|\\
&\le 2\left\|\sqrt{n}\left(\widehat{\theta}_n-\theta_0\right)\right\|.\frac{1}{n}\sum_{t=1}^{n}\sup_{||\theta-\theta_0||\le||\widehat{\theta}_n-\theta_0||}\|\nabla f_{t-1}(\theta)-\nabla f_{t-1}(\theta_0)\|\\
&+\sup_{\tau\in\Upsilon}|\tau|\times\left\|\sqrt{n}\left(\widehat{\theta}_n-\theta_0\right)\right\|\\
&\qquad\times\sum_{j=1}^{k}\frac{1}{n}\sum_{t=1}^{n}||\nabla f_{t-1}(\theta_0)||.\left|\Phi\left(\underline{\widetilde{Y}}_{t-j}\right)-\Phi\left(\underline{Y}_{t-j}\right)\right|
\end{aligned}$$

Theorem 2.2 implies that $\left\|\sqrt{n}\left(\widehat{\theta}_n-\theta_0\right)\right\|=O_p(1)$, and it follows from Lemma 8.1 that for fixed k, $\sum_{j=1}^{k}\frac{1}{n}\sum_{t=1}^{n}||\nabla f_{t-1}(\theta_0)||.\left|\Phi\left(\underline{\widetilde{Y}}_{t-j}\right)-\Phi\left(\underline{Y}_{t-j}\right)\right|=o_p(1)$.

Moreover, part (2.4) of Assumption 2.2 implies that for an arbitrary small $\varepsilon > 0$, $\kappa(\varepsilon) = E\left[\sup_{||\theta-\theta_0||\leq\varepsilon} \|\nabla f_{t-1}(\theta) - \nabla f_{t-1}(\theta_0)\|\right] < \infty$, so that by Lemma 2.2,

$$p\lim_{n\to\infty} \frac{1}{n}\sum_{t=1}^{n} \sup_{||\theta-\theta_0||\leq\varepsilon} \|\nabla f_{t-1}(\theta) - \nabla f_{t-1}(\theta_0)\| = \kappa(\varepsilon).$$

It follows now easily that

$$\begin{aligned} p\lim_{n\to\infty} \frac{1}{n}\sum_{t=1}^{n} \sup_{||\theta-\theta_0||\leq||\widehat{\theta}_n-\theta_0||} \|\nabla f_{t-1}(\theta) - \nabla f_{t-1}(\theta_0)\| &= p\lim_{n\to\infty} \kappa\left(||\widehat{\theta}_n - \theta_0||\right) \\ &= \kappa(0) = 0. \end{aligned} \tag{8.14}$$

Thus, $\sup_{(\tau_1,\tau_2,....\tau_k)'\in\Upsilon^k} \left|\mathrm{Re}\left[\widetilde{W}_{k,n}^{(2)}(\tau_1,\tau_2,....\tau_k)\right] - \mathrm{Re}\left[\widehat{W}_{k,n}^{(2)}(\tau_1,\tau_2,....\tau_k)\right]\right| = o_p(1)$, and the same applies to the Im[.] case. ■

Next, denote

$$\begin{aligned} \widehat{B}_{n,k} &= \int_{\Upsilon^k} \left|\widehat{W}_{k,n}(\tau_1,\tau_2,....\tau_k)\right|^2 \mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k), \\ \widetilde{B}_{n,k} &= \int_{\Upsilon^k} \left|\widetilde{W}_{k,n}(\tau_1,\tau_2,....\tau_k)\right|^2 \mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k). \end{aligned}$$

Then

Lemma 8.3. *Under H_0 and Assumptions 1.1, 2.1 and 2.2, $\widetilde{B}_{n,k} = \widehat{B}_{n,k} + o_p(1)$ for each $k \in \mathbb{N}$.*

Proof. Lemma 8.2 implies that

$$\widetilde{W}_{k,n}(\tau_1,\tau_2,....\tau_k) - \widehat{W}_{k,n}(\tau_1,\tau_2,....\tau_k) = \Delta W_{k,n}(\tau_1,\tau_2,....\tau_k) + o_p(1),$$

uniformly in $(\tau_1,\tau_2,...,\tau_k)' \in \Upsilon^k$, where

$$\begin{aligned} &\Delta W_{k,n}(\tau_1,\tau_2,....\tau_k) \\ &\quad = \frac{1}{\sqrt{n}}\sum_{t=1}^{n} U_t\left(\exp\left(\mathbf{i}.\sum_{j=1}^{k}\tau_j\Phi\left(\underline{\widetilde{Y}}_{t-j}\right)\right) - \exp\left(\mathbf{i}.\sum_{j=1}^{k}\tau_j\Phi\left(\underline{Y}_{t-j}\right)\right)\right). \end{aligned}$$

Hence

$$
\begin{aligned}
\widetilde{B}_{n,k} &= \int_{\Upsilon^k} \left| \widehat{W}_{k,n}(\tau_1, \tau_2, ...\tau_k) + \Delta W_{k,n}(\tau_1, \tau_2, ...\tau_k) + o_p(1) \right|^2 \\
&\quad \times \mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k) \\
&= \int_{\Upsilon^k} \left(\mathrm{Re}\left[\widehat{W}_{k,n}(\tau_1, \tau_2, ...\tau_k) \right] \right)^2 \mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k) \\
&\quad + \int_{\Upsilon^k} \left(\mathrm{Im}\left[\widehat{W}_{k,n}(\tau_1, \tau_2, ...\tau_k) \right] \right)^2 \mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k) \\
&\quad + \int_{\Upsilon^k} \left(\mathrm{Re}\left[\Delta W_{k,n}(\tau_1, \tau_2, ...\tau_k) \right] \right)^2 \mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k) \\
&\quad + \int_{\Upsilon^k} \left(\mathrm{Im}\left[\Delta W_{k,n}(\tau_1, \tau_2, ...\tau_k) \right] \right)^2 \mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k) \\
&\quad + 2 \int_{\Upsilon^k} \mathrm{Re}\left[\widehat{W}_{k,n}(\tau_1, \tau_2, ...\tau_k) \right] \left(\mathrm{Re}\left[\Delta W_{k,n}(\tau_1, \tau_2, ...\tau_k) \right] \right) \\
&\quad \times \mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k) \\
&\quad + 2 \int_{\Upsilon^k} \mathrm{Im}\left[\widehat{W}_{k,n}(\tau_1, \tau_2, ...\tau_k) \right] \left(\mathrm{Im}\left[\Delta W_{k,n}(\tau_1, \tau_2, ...\tau_k) \right] \right) \\
&\quad \times \mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k) \\
&\quad + 2 \int_{\Upsilon^k} \mathrm{Re}\left[\widehat{W}_{k,n}(\tau_1, \tau_2, ...\tau_k) \right] \mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k) \times o_p(1) \\
&\quad + 2 \int_{\Upsilon^k} \mathrm{Im}\left[\widehat{W}_{k,n}(\tau_1, \tau_2, ...\tau_k) \right] \mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k) \times o_p(1) \\
&\quad + 2 \int_{\Upsilon^k} \mathrm{Re}\left[\Delta W_{k,n}(\tau_1, \tau_2, ...\tau_k) \right] \mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k) \times o_p(1) \\
&\quad + 2 \int_{\Upsilon^k} \mathrm{Im}\left[\Delta W_{k,n}(\tau_1, \tau_2, ...\tau_k) \right] \mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k) \times o_p(1) \\
&\quad + o_p(1)
\end{aligned}
$$

and thus,

$$
\begin{aligned}
&\left| \widetilde{B}_{n,k} - \widehat{B}_{n,k} \right| \\
&\leq \int_{\Upsilon^k} \left(\mathrm{Re}\left[\Delta W_{k,n}(\tau_1, \tau_2, ...\tau_k) \right] \right)^2 \mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k) \\
&+ \int_{\Upsilon^k} \left(\mathrm{Im}\left[\Delta W_{k,n}(\tau_1, \tau_2, ...\tau_k) \right] \right)^2 \mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k)
\end{aligned}
$$

$$
\begin{aligned}
&+2\sqrt{\int_{\Upsilon^k}\left(\mathrm{Re}\left[\widehat{W}_{k,n}(\tau_1,\tau_2,....\tau_k)\right]\right)^2\mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k)}\\
&\qquad\times\sqrt{\int_{\Upsilon^k}\left(\mathrm{Re}\left[\Delta W_{k,n}(\tau_1,\tau_2,....\tau_k)\right]\right)^2\mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k)}\\
&+2\sqrt{\int_{\Upsilon^k}\left(\mathrm{Im}\left[\widehat{W}_{k,n}(\tau_1,\tau_2,....\tau_k)\right]\right)^2\mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k)}\\
&\qquad\times\sqrt{\int_{\Upsilon^k}\left(\mathrm{Im}\left[\Delta W_{k,n}(\tau_1,\tau_2,....\tau_k)\right]\right)^2\mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k)}\\
&+2\sqrt{\int_{\Upsilon^k}\left(\mathrm{Re}\left[\widehat{W}_{k,n}(\tau_1,\tau_2,....\tau_k)\right]\right)^2\mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k)}\times o_p(1)\\
&+2\sqrt{\int_{\Upsilon^k}\left(\mathrm{Im}\left[\widehat{W}_{k,n}(\tau_1,\tau_2,....\tau_k)\right]\right)^2\mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k)}\times o_p(1)\\
&+2\sqrt{\int_{\Upsilon^k}\left(\mathrm{Re}\left[\Delta W_{k,n}(\tau_1,\tau_2,....\tau_k)\right]\right)^2\mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k)}\times o_p(1)\\
&+2\int_{\Upsilon^k}\left(\mathrm{Im}\left[\Delta W_{k,n}(\tau_1,\tau_2,....\tau_k)\right]\right)^2\mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k)\times o_p(1)\\
&+o_p(1)\\
&\le\int_{\Upsilon^k}\left|\Delta W_{k,n}(\tau_1,\tau_2,....\tau_k)\right|^2\mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k)\\
&+4\sqrt{\widehat{B}_{n,k}}\times\int_{\Upsilon^k}\left|\Delta W_{k,n}(\tau_1,\tau_2,....\tau_k)\right|^2\mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k)\\
&+4\sqrt{\widehat{B}_{n,k}}\times o_p(1)\\
&+4\sqrt{\int_{\Upsilon^k}\left|\Delta W_{k,n}(\tau_1,\tau_2,....\tau_k)\right|^2\mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k)}\times o_p(1)\\
&+o_p(1)
\end{aligned}
$$

Since by Lemma 3.5, $\widehat{B}_{n,k}\xrightarrow{\mathrm{d}}B_k$, which implies $\widehat{B}_{n,k}=O_p(1)$, it suffices to prove that

$$\int_{\Upsilon^k}\left|\Delta W_{k,n}(\tau_1,\tau_2,....\tau_k)\right|^2\mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k)=o_p(1), \tag{8.15}$$

as follows. Note that for $(\tau_1, \tau_2, ...\tau_k)' \in \Upsilon^k$,

$$
\begin{aligned}
& E\left[\left|\Delta W_{k,n}(\tau_1, \tau_2, ...\tau_k)\right|^2\right] \\
= \; & \frac{1}{n}\sum_{t=1}^{n} E\left[U_t^2\left(\cos\left(\sum_{j=1}^{k}\tau_j\Phi\left(\widetilde{\underline{Y}}_{t-j}\right)\right) - \cos\left(\sum_{j=1}^{k}\tau_j\Phi\left(\underline{Y}_{t-j}\right)\right)\right)^2\right] \\
& + \frac{1}{n}\sum_{t=1}^{n} E\left[U_t^2\left(\sin\left(\sum_{j=1}^{k}\tau_j\Phi\left(\widetilde{\underline{Y}}_{t-j}\right)\right) - \sin\left(\sum_{j=1}^{k}\tau_j\Phi\left(\underline{Y}_{t-j}\right)\right)\right)^2\right] \\
= \; & 2E[U_t^2] \\
& - 2\frac{1}{n}\sum_{t=1}^{n} E\left[U_t^2\cos\left(\sum_{j=1}^{k}\tau_j\Phi\left(\widetilde{\underline{Y}}_{t-j}\right)\right) \times \cos\left(\sum_{j=1}^{k}\tau_j\Phi\left(\underline{Y}_{t-j}\right)\right)\right] \\
& - 2\frac{1}{n}\sum_{t=1}^{n} E\left[U_t^2\sin\left(\sum_{j=1}^{k}\tau_j\Phi\left(\widetilde{\underline{Y}}_{t-j}\right)\right) \times \sin\left(\sum_{j=1}^{k}\tau_j\Phi\left(\underline{Y}_{t-j}\right)\right)\right] \\
= \; & 2E[U_1^2] - 2\frac{1}{n}\sum_{t=1}^{n} E\left[U_t^2\cos\left(\sum_{j=1}^{k}\tau_j\left(\Phi\left(\widetilde{\underline{Y}}_{t-j}\right) - \Phi\left(\underline{Y}_{t-j}\right)\right)\right)\right] \\
= \; & 2\frac{1}{n}\sum_{t=1}^{n} E\left[U_t^2\left(1 - \cos\left(\sum_{j=1}^{k}\tau_j\left(\Phi\left(\widetilde{\underline{Y}}_{t-j}\right) - \Phi\left(\underline{Y}_{t-j}\right)\right)\right)\right)\right] \\
\leq \; & 2\sum_{j=1}^{k}|\tau_j.|\frac{1}{n}\sum_{t=1}^{n} E\left[U_t^2\left|\Phi\left(\widetilde{\underline{Y}}_{t-j}\right) - \Phi\left(\underline{Y}_{t-j}\right)\right|\right] \\
\leq \; & 2\sup_{\tau\in\Upsilon}|\tau|\sum_{j=1}^{k}\frac{1}{n}\sum_{t=1}^{n} E\left[U_t^2\left|\Phi\left(\widetilde{\underline{Y}}_{t-j}\right) - \Phi\left(\underline{Y}_{t-j}\right)\right|\right] \\
= \; & o(1)
\end{aligned}
$$

where the first inequality follows from the mean value theorem and the last equality is due to Lemma 8.1. Thus,

$$
\int_{\Upsilon^k} E\left[\left|\Delta W_{k,n}(\tau_1, \tau_2, ...\tau_k)\right|^2\right] \mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k) = o(1)
$$

which implies (8.15). ■

It follows now from Theorem 3.1, Lemma 3.1 and Lemma 8.3 that the following general result holds.

Theorem 8.1. *Denote*

$$\begin{aligned}\widetilde{W}_{k,n}(\tau_1,\tau_2,....\tau_k) &= \frac{1}{\sqrt{n}}\sum_{t=1}^{n}(Y_t - f_{t-1}(\widehat{\theta}_n))\exp\left(\mathbf{i}.\sum_{j=1}^{k}\tau_j'\Phi(\underline{\widetilde{Z}}_{t-j})\right),\\ \widetilde{B}_{n,k} &= \int_{\Upsilon^k}\left|\widetilde{W}_{k,n}(\tau_1,\tau_2,....\tau_k)\right|^2 \mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k),\\ \widetilde{T}_n &= \sum_{k=1}^{L_n}\gamma_k\widetilde{B}_{n,k},\end{aligned}$$

where each component $\underline{\widetilde{Z}}_{i,t}$ of $\underline{\widetilde{Z}}_t$ is a standardized version of component $Z_{i,t}$ of Z_t, similar to $\underline{\widetilde{Y}}_t$ in (8.6) *with Y_t replaced by $Z_{i,t}$. Similarly, denote*

$$\begin{aligned}\widehat{W}_{k,n}(\tau_1,\tau_2,....\tau_k) &= \frac{1}{\sqrt{n}}\sum_{t=1}^{n}(Y_t - f_{t-1}(\widehat{\theta}_n))\exp\left(\mathbf{i}.\sum_{j=1}^{k}\tau_j'\Phi(\underline{Z}_{t-j})\right),\\ \widehat{B}_{n,k} &= \int_{\Upsilon^k}\left|\widehat{W}_{k,n}(\tau_1,\tau_2,....\tau_k)\right|^2 \mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k),\\ \widehat{T}_n &= \sum_{k=1}^{L_n}\gamma_k\widehat{B}_{n,k}.\end{aligned}$$

where each component $\underline{Z}_{i,t}$ of $\underline{Z}_t$ is a standardized version of component $Z_{i,t}$ of Z_t, similar to $\underline{Y}_t$ in (8.6) *with Y_t replaced by $Z_{i,t}$. Moreover, let all the components of the one-to one mapping Φ be* arctan(.) *functions. Then under the conditions of Theorem 3.1,*

$$\widetilde{T}_n = \widehat{T}_n + o_p(1) \xrightarrow{\mathrm{d}} \sum_{k=1}^{\infty}\gamma_k B_k = T,$$

where for each k the random variable B_k represents the limiting distribution of $\widehat{B}_{n,k}$. Therefore, all the previous result for the WICM test statistic $\widehat{T}_n$ under H_0 carry over to $\widetilde{T}_n$.

Proof. Recall from Theorem 3.1 that

$$\widehat{T}_n = \sum_{k=1}^{L_n}\gamma_k\widehat{B}_{n,k} \xrightarrow{\mathrm{d}} \sum_{k=1}^{\infty}\gamma_k B_k = T \tag{8.16}$$

and note that Lemma 3.1 carries over to $\widetilde{B}_{n,k}$, i.e., $\widetilde{B}_{n,k} \leq 2.\widetilde{B}_{1,n,k} + 2\widetilde{B}_{2,n,k}$, where $E[\widetilde{B}_{1,n,k}] = E[U_1^2]$ and $\sup_{k\in\mathbb{N}} \widetilde{B}_{2,n,k} = O_p(1)$. Of course, Lemma 3.1 is directly applicable to the current $\widehat{B}_{n,k}$'s as well. Then with $\widehat{B}_{1,n,k}$ and $\widehat{B}_{2,n,k}$ as in Lemma 3.1, it follows that for each $K \in \mathbb{N}$,

$$\begin{aligned}
\sum_{k=K+1}^{\infty} \gamma_k \left|\widetilde{B}_{n,k} - \widehat{B}_{n,k}\right| &\leq 2\sum_{k=K+1}^{\infty} \gamma_k \left(\widetilde{B}_{1,n,k} + \widehat{B}_{1,n,k}\right) \\
&\quad +2\sum_{k=K+1}^{\infty} \gamma_k \sup_{k\in\mathbb{N}} \left(\widetilde{B}_{2,n,k} + \widehat{B}_{2,n,k}\right) \\
&= \sum_{k=K+1}^{\infty} \gamma_k \times O_p(1) \qquad (8.17)
\end{aligned}$$

whereas by Lemma 8.3,

$$\sum_{k=1}^{K} \gamma_k \left|\widetilde{B}_{n,k} - \widehat{B}_{n,k}\right| = o_p(1). \qquad (8.18)$$

The result (8.17) implies that for arbitrary $\varepsilon \in (0,1)$ and $\delta > 0$ we can choose K so large that

$$\limsup_{n\to\infty} \Pr\left[\sum_{k=K+1}^{\infty} \gamma_k \left|\widetilde{B}_{n,k} - \widehat{B}_{n,k}\right| > \delta/2\right] < \varepsilon/2,$$

whereas (8.18) implies that for this K,

$$\limsup_{n\to\infty} \Pr\left[\sum_{k=1}^{K} \gamma_k \left|\widetilde{B}_{n,k} - \widehat{B}_{n,k}\right| > \delta/2\right] < \varepsilon/2.$$

It follows now from the inequality (3.15) that

$$\limsup_{n\to\infty} \Pr\left[\sum_{k=1}^{\infty} \gamma_k \left|\widetilde{B}_{n,k} - \widehat{B}_{n,k}\right| > \delta\right] < \varepsilon,$$

hence

$$\sum_{k=1}^{L_n} \gamma_k \left|\widetilde{B}_{n,k} - \widehat{B}_{n,k}\right| \leq \sum_{k=1}^{\infty} \gamma_k \left|\widetilde{B}_{n,k} - \widehat{B}_{n,k}\right| = o_p(1). \qquad (8.19)$$

The theorem under review now follows from (8.16) and (8.19). ■

8.2.2. The case H_1

Also the consistency result in Theorem 4.1 carries over to the WICM test statistic $\widetilde{T}_n$.

Theorem 8.2. *Under H_1 and Assumptions 3.1 and 4.1, and under the notations in Theorems 4.1 and 8.1,*

$$\widetilde{T}_n/n = \sum_{k=1}^{L_n} \gamma_k \widetilde{B}_{k,n}/n = \widehat{T}_n/n + o_p(1) \xrightarrow{\mathrm{p}} \sum_{k=1}^{\infty} \gamma_k \eta_k > 0.$$

Proof. Denote

$$\begin{aligned}
\widehat{B}_{n,k}^{(1)} &= \int_{\Upsilon^k} \left|\widehat{W}_{k,n}^{(1)}(\tau_1, \tau_2, ...\tau_k)\right|^2 \mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k), \\
\widetilde{B}_{n,k}^{(1)} &= \int_{\Upsilon^k} \left|\widetilde{W}_{k,n}^{(1)}(\tau_1, \tau_2, ...\tau_k)\right|^2 \mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k),
\end{aligned}$$

$$\begin{aligned}
\widehat{B}_{n,k}^{(2)} &= \int_{\Upsilon^k} \left|\widehat{W}_{k,n}^{(2)}(\tau_1, \tau_2, ...\tau_k)\right|^2 \mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k), \\
\widetilde{B}_{n,k}^{(2)} &= \int_{\Upsilon^k} \left|\widetilde{W}_{k,n}^{(2}(\tau_1, \tau_2, ...\tau_k)\right|^2 \mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k),
\end{aligned}$$

and let $\widehat{B}_{n,k}$ and $\widetilde{B}_{n,k}$ be the same as in Theorem 8.1. Moreover, denote

$$\begin{aligned}
\psi_k(\tau_1, \tau_2, ...\tau_k) &= E\left[U_t \exp\left(\mathbf{i}.\sum_{j=1}^{k} \tau_j \Phi\left(\underline{Y}_{t-j}\right)\right)\right] \\
\eta_k &= \int_{\Upsilon^k} \left|\psi_k(\tau_1, \tau_2, ...\tau_k)\right|^2 \mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k)
\end{aligned}$$

I will prove the theorem under review by showing that

$$\sup_{k\in\mathbb{N}} \widetilde{B}_{n,k}^{(1)} = O_p(1),\ \sup_{k\in\mathbb{N}} \widehat{B}_{n,k}^{(1)} = O_p(1), \tag{8.20}$$

$$\sup_{k\in\mathbb{N}} \widetilde{B}_{n,k}^{(2)} = o_p(1),\ \sup_{k\in\mathbb{N}} \widehat{B}_{n,k}^{(2)} = o_p(1), \tag{8.21}$$

$$\sup_{k\in\mathbb{N}} \left|\widetilde{B}_{n,k}/n - \widetilde{B}_{n,k}^{(1)}/n\right| = o_p(1),\ \sup_{k\in\mathbb{N}} \left|\widehat{B}_{n,k}/n - \widehat{B}_{n,k}^{(1)}/n\right| = o_p(1), \tag{8.22}$$

$$\left|\widetilde{B}_{n,k}^{(1)}/n - \widehat{B}_{n,k}^{(1)}/n\right| = o_p(1),\ \text{for each } k \in \mathbb{N}, \tag{8.23}$$

$$\widehat{B}_{n,k}^{(1)}/n \xrightarrow{\mathrm{p}} \eta_k,\ \text{for each } k \in \mathbb{N}. \tag{8.24}$$

Then it not hard to verify that Theorem 8.2 holds.

Proof of (8.20). It is easy to verify from (8.10) and (8.12) that

$$\begin{array}{l}\sup_{k\in\mathbb{N}}\sup_{(\tau_1,\tau_2,....\tau_k)'\in\Upsilon^k}\left|n^{-1/2}\widehat{W}^{(1)}_{k,n}(\tau_1,\tau_2,....\tau_k)\right|\leq C_{1,n},\\ \sup_{k\in\mathbb{N}}\sup_{(\tau_1,\tau_2,....\tau_k)'\in\Upsilon^k}\left|n^{-1/2}\widetilde{W}^{(1)}_{k,n}(\tau_1,\tau_2,....\tau_k)\right|\leq C_{1,n},\end{array}$$

where

$$C_{1,n}=\sqrt{2}\frac{1}{n}\sum_{t=1}^{n}|U_t|=O_p(1).$$

Hence,

$$\sup_{k\in\mathbb{N}}\widetilde{B}^{(1)}_{n,k}\leq C^2_{1,n}=O_p(1),\ \sup_{k\in\mathbb{N}}\widehat{B}^{(1)}_{n,k}\leq C^2_{1,n}=O_p(1). \tag{8.25}$$

Proof of (8.21). Again, it is easy to verify from (8.11) and (8.13) that

$$\begin{array}{l}\sup_{k\in\mathbb{N}}\sup_{(\tau_1,\tau_2,....\tau_k)'\in\Upsilon^k}\left|n^{-1/2}\widehat{W}^{(2)}_{k,n}(\tau_1,\tau_2,....\tau_k)\right|\leq C_{2,n},\\ \sup_{k\in\mathbb{N}}\sup_{(\tau_1,\tau_2,....\tau_k)'\in\Upsilon^k}\left|n^{-1/2}\widetilde{W}^{(2)}_{k,n}(\tau_1,\tau_2,....\tau_k)\right|\leq C_{2,n},\end{array}$$

where

$$\begin{aligned}C_{2,n} &= \sqrt{2}\frac{1}{n}\sum_{t=1}^{n}\left|f_{t-1}(\widehat{\theta}_n)-f_{t-1}(\theta_0)\right|\\ &\leq \sqrt{2}\frac{1}{n}\sum_{t=1}^{n}\sup_{||\theta-\theta_0||\leq||\widehat{\theta}_n-\theta_0||}|f_{t-1}(\theta)-f_{t-1}(\theta_0)|\\ &= o_p(1).\end{aligned}$$

The latter follows from $||\widehat{\theta}_n-\theta_0||\xrightarrow{\mathrm{p}}0$ and (8.14). Hence

$$\sup_{k\in\mathbb{N}}\widetilde{B}^{(2)}_{n,k}\leq C^2_{2,n}=o_p(1),\ \sup_{k\in\mathbb{N}}\widehat{B}^{(2)}_{n,k}\leq C^2_{2,n}=o_p(1). \tag{8.26}$$

Proof of (8.22). Observe from (8.3) that

$$\begin{aligned}\widetilde{B}_{n,k}/n &= \int_{\Upsilon^k}\left|n^{-1/2}\widetilde{W}_{k,n}(\tau_1,\tau_2,....\tau_k)\right|^2\mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k)\\ &= \int_{\Upsilon^k}\left(\mathrm{Re}\left[n^{-1/2}\widetilde{W}^{(1)}_{k,n}(\tau_1,\tau_2,....\tau_k)\right]-\mathrm{Re}\left[n^{-1/2}\widetilde{W}^{(2)}_{k,n}(\tau_1,\tau_2,....\tau_k)\right]\right)^2\end{aligned}$$

$$
\begin{aligned}
&\times \mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k)\\
&+\int_{\Upsilon^k}\left(\mathrm{Im}\left[n^{-1/2}\widetilde{W}^{(1)}_{k,n}(\tau_1,\tau_2,....\tau_k)\right]-\mathrm{Im}\left[n^{-1/2}\widetilde{W}^{(2)}_{k,n}(\tau_1,\tau_2,....\tau_k)\right]\right)^2\\
&\times \mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k)\\
=\;& \widetilde{B}^{(1)}_{n,k}/n+\int_{\Upsilon^k}\left|n^{-1/2}\widetilde{W}^{(2)}_{k,n}(\tau_1,\tau_2,....\tau_k)\right|^2\mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k)\\
&-2\int_{\Upsilon^k}\mathrm{Re}\left[n^{-1/2}\widetilde{W}^{(1)}_{k,n}(\tau_1,\tau_2,....\tau_k)\right]\times\mathrm{Re}\left[n^{-1/2}\widetilde{W}^{(2)}_{k,n}(\tau_1,\tau_2,....\tau_k)\right]\\
&\times \mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k)\\
&-2\int_{\Upsilon^k}\mathrm{Im}\left[n^{-1/2}\widetilde{W}^{(1)}_{k,n}(\tau_1,\tau_2,....\tau_k)\right]\times\mathrm{Im}\left[n^{-1/2}\widetilde{W}^{(2)}_{k,n}(\tau_1,\tau_2,....\tau_k)\right]\\
&\times \mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k),
\end{aligned}
$$

hence

$$
\begin{aligned}
&\left|\widetilde{B}_{n,k}/n-\widetilde{B}^{(1)}_{n,k}/n\right|\le \widetilde{B}^{(2)}_{n,k}/n\\
&+2\int_{\Upsilon^k}\left|\mathrm{Re}\left[n^{-1/2}\widetilde{W}^{(1)}_{k,n}(\tau_1,\tau_2,....\tau_k)\right]\right|\times\left|\mathrm{Re}\left[n^{-1/2}\widetilde{W}^{(2)}_{k,n}(\tau_1,\tau_2,....\tau_k)\right]\right|\\
&\times \mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k)\\
&+2\int_{\Upsilon^k}\left|\mathrm{Im}\left[n^{-1/2}\widetilde{W}^{(1)}_{k,n}(\tau_1,\tau_2,....\tau_k)\right]\right|\times\left|\mathrm{Im}\left[n^{-1/2}\widetilde{W}^{(2)}_{k,n}(\tau_1,\tau_2,....\tau_k)\right]\right|\\
&\times \mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k)\\
&\le \widetilde{B}^{(2)}_{n,k}/n+4\sqrt{\widetilde{B}^{(1)}_{n,k}/n}\sqrt{\widetilde{B}^{(2)}_{n,k}/n}\\
&\le C^2_{2,n}+4.C_{1,n}.C_{2,n}=o_p(1),
\end{aligned}
$$

where the latter follows from (3.14) and (8.26). Obviously, the last inequality also holds for $\widehat{B}_{n,k}/n-\widehat{B}^{(1)}_{n,k}/n$. This proves (8.22).

Proof of (8.23). Denote

$$
\Delta W^{(1)}_{k,n}(\tau_1,\tau_2,....\tau_k)=\widetilde{W}^{(1)}_{k,n}(\tau_1,\tau_2,....\tau_k)-\widehat{W}^{(1)}_{k,n}(\tau_1,\tau_2,....\tau_k).
$$

and

$$
B^{\Delta}_{k,n}=\int_{\Upsilon^k}\left|\Delta W^{(1)}_{k,n}(\tau_1,\tau_2,....\tau_k)\right|^2\mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k).
$$

It is not hard to verify from (8.10) and (8.12) and the mean value theorem that

$$
\sup_{(\tau_1,\tau_2,....\tau_k)'\in\Upsilon^k}\left|n^{-1/2}\Delta W^{(1)}_{k,n}(\tau_1,\tau_2,....\tau_k)\right|\le D_{k,n},
$$

where

$$D_{k,n} = \sqrt{2} \sup_{\tau \in \Upsilon} |\tau| \sum_{j=1}^{k} \frac{1}{n} \sum_{t=1}^{n} |U_t|. \left| \Phi \left(\underline{\widetilde{Y}}_{t-j} \right) - \Phi \left(\underline{Y}_{t-j} \right) \right| = o_p(1),$$

where the latter follows from Lemma 8.1. Hence

$$B_{k,n}^{\Delta}/n \leq D_{k,n}^2 = o_p(1). \tag{8.27}$$

Then

$$\begin{aligned}
&\widetilde{B}_{n,k}^{(1)}/n \\
&= \int_{\Upsilon^k} \left| n^{-1/2} \widehat{W}_{k,n}^{(1)}(\tau_1, \tau_2, ...\tau_k) + n^{-1/2} \Delta W_{k,n}^{(1)}(\tau_1, \tau_2, ...\tau_k) \right|^2 \\
&\times \mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k) \\
&= \int_{\Upsilon^k} \left(\mathrm{Re} \left[n^{-1/2} \widehat{W}_{k,n}^{(1)}(\tau_1, \tau_2, ...\tau_k) \right] + \mathrm{Re} \left[n^{-1/2} \Delta W_{k,n}^{(1)}(\tau_1, \tau_2, ...\tau_k) \right] \right)^2 \\
&\times \mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k) \\
&+ \int_{\Upsilon^k} \left(\mathrm{Im} \left[n^{-1/2} \widehat{W}_{k,n}^{(1)}(\tau_1, \tau_2, ...\tau_k) \right] + \mathrm{Im} \left[n^{-1/2} \Delta W_{k,n}^{(1)}(\tau_1, \tau_2, ...\tau_k) \right] \right)^2 \\
&\times \mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k) \\
&= \int_{\Upsilon^k} \left(\mathrm{Re} \left[n^{-1/2} \widehat{W}_{k,n}^{(1)}(\tau_1, \tau_2, ...\tau_k) \right] \right)^2 \mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k) \\
&+ \int_{\Upsilon^k} \left(\mathrm{Im} \left[n^{-1/2} \widehat{W}_{k,n}^{(1)}(\tau_1, \tau_2, ...\tau_k) \right] \right)^2 \mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k) \\
&+ \int_{\Upsilon^k} \left(\mathrm{Re} \left[n^{-1/2} \Delta W_{k,n}^{(1)}(\tau_1, \tau_2, ...\tau_k) \right] \right)^2 \mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k) \\
&+ \int_{\Upsilon^k} \left(\mathrm{Im} \left[n^{-1/2} \Delta W_{k,n}^{(1)}(\tau_1, \tau_2, ...\tau_k) \right] \right) \mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k) \\
&+2 \int_{\Upsilon^k} \left(\mathrm{Re} \left[n^{-1/2} \widehat{W}_{k,n}^{(1)}(\tau_1, \tau_2, ...\tau_k) \right] \right) \left(\mathrm{Re} \left[n^{-1/2} \Delta W_{k,n}^{(1)}(\tau_1, \tau_2, ...\tau_k) \right] \right) \\
&\times \mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k) \\
&+2 \int_{\Upsilon^k} \left(\mathrm{Im} \left[n^{-1/2} \widehat{W}_{k,n}^{(1)}(\tau_1, \tau_2, ...\tau_k) \right] \right) \left(\mathrm{Im} \left[n^{-1/2} \Delta W_{k,n}^{(1)}(\tau_1, \tau_2, ...\tau_k) \right] \right) \\
&\times \mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k) \\
&= \widehat{B}_{n,k}^{(1)}/n + B_{k,n}^{\Delta}/n
\end{aligned}$$

$$+2\int_{\Upsilon^k}\left(\mathrm{Re}\left[n^{-1/2}\widehat{W}_{k,n}^{(1)}(\tau_1,\tau_2,....\tau_k)\right]\right)\left(\mathrm{Re}\left[n^{-1/2}\Delta W_{k,n}^{(1)}(\tau_1,\tau_2,....\tau_k)\right]\right)$$
$$\times\mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k)$$
$$+2\int_{\Upsilon^k}\left(\mathrm{Im}\left[n^{-1/2}\widehat{W}_{k,n}^{(1)}(\tau_1,\tau_2,....\tau_k)\right]\right)\left(\mathrm{Im}\left[n^{-1/2}\Delta W_{k,n}^{(1)}(\tau_1,\tau_2,....\tau_k)\right]\right)$$
$$\times\mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k),$$

hence

$$\begin{aligned}
&\left|\widetilde{B}_{n,k}^{(1)}/n-\widehat{B}_{n,k}^{(1)}/n\right|\leq B_{k,n}^{\Delta}/n\\
&+2\sqrt{\int_{\Upsilon^k}\left(\mathrm{Re}\left[n^{-1/2}\widehat{W}_{k,n}^{(1)}(\tau_1,\tau_2,....\tau_k)\right]\right)^2\mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k)}\\
&\overline{\times\int_{\Upsilon^k}\left(\mathrm{Re}\left[n^{-1/2}\Delta W_{k,n}^{(1)}(\tau_1,\tau_2,....\tau_k)\right]\right)^2\mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k)}\\
&+2\sqrt{\int_{\Upsilon^k}\left(\mathrm{Im}\left[n^{-1/2}\widehat{W}_{k,n}^{(1)}(\tau_1,\tau_2,....\tau_k)\right]\right)^2\mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k)}\\
&\overline{\times\int_{\Upsilon^k}\left(\mathrm{Im}\left[n^{-1/2}\Delta W_{k,n}^{(1)}(\tau_1,\tau_2,....\tau_k)\right]\right)^2\mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k)}\\
&\leq B_{k,n}^{\Delta}/n+4\sqrt{\widehat{B}_{n,k}^{(1)}/n}\sqrt{B_{k,n}^{\Delta}/n}\leq D_{k,n}^2+4C_{1,n}D_{k,n}=o_p(1),
\end{aligned}$$

where the latter follows from (3.14) and (8.27).

Proof of (8.26). Next, denote

$$\psi_k(\tau_1,\tau_2,....\tau_k)=E\left[U_t\exp\left(\mathbf{i}.\sum_{j=1}^{k}\tau_j\Phi\left(\underline{Y}_{t-j}\right)\right)\right],$$
$$\eta_k=\int_{\Upsilon^k}\left|\psi_k(\tau_1,\tau_2,....\tau_k)\right|^2\mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k),$$

and observe from Lemma 2.1 that under Assumptions 1.1 and 2.1,

$$\sup_{(\tau_1,\tau_2,....\tau_k)'\in\Upsilon^k}\left|n^{-1/2}\widehat{W}_{k,n}^{(1)}(\tau_1,\tau_2,....\tau_k)-\psi_k(\tau_1,\tau_2,....\tau_k)\right|=o_p(1)$$

for each $k\in\mathbb{N}$, which implies that $\widehat{B}_{n,k}^{(1)}/n\xrightarrow{\mathrm{p}}\eta_k$ for each $k\in\mathbb{N}$. ■

9. Models with infinitely many lagged conditioning variables

So far the above results have been derived for nonlinear ARX models for which the conditional expectation function $f_{t-1}(\theta_0)$ depend on a finite number of lagged dependent variables Y_t and a finite number of lagged exogenous variables X_t. However, since $f_{t-1}(\theta_0)$ aims to represent the conditional expectation $E[Y_t|\mathcal{F}_{-\infty}^{t-1}]$, which in general may depend on all lagged Y_t's and lagged X_t's, it may be more realistic to allow infinitely many lags in the specification of $f_{t-1}(\theta_0)$. A convenient way to do that is to specify the errors U_t as an ARMA process, for example an ARMA(1,1) process $U_t = \kappa_0 U_{t-1} + e_t - \delta_0 e_{t-1}$, where $|\kappa_0| < 1$, $|\delta_0| < 1$ and $\delta_0 \neq \kappa_0$, and now under H_0, $E[e_t|\mathcal{F}_{-\infty}^{t-1}] = 0$ a.s. Then model (1.2) becomes a nonlinear ARMAX model,

$$Y_t = f_{t-1}(\theta_0) + \kappa_0(Y_{t-1} - f_{t-2}(\theta_0)) + e_t - \delta_0 e_{t-1},$$

which by inverting the lag polynomial $1 - \delta_0 L$ can be written as an infinite-order ARX model,

$$\begin{aligned} Y_t &= f_{t-1}(\theta_0) + (\kappa_0 - \delta_0)\sum_{j=0}^{\infty}\delta_0^j(Y_{t-1-j} - f_{t-2-j}(\theta_0)) + e_t \\ &= g_{t-1}(\theta_0, \delta_0, \kappa_0) + e_t, \end{aligned}$$

say.

Previously it was assumed that $Z_t = (Y_t, X_t')'$ is observed for $1 - \max(p, q) \leq t \leq n$, where p is the maximum lag of Y_t in $f_{t-1}(\theta_0)$, and q is the maximum lag of X_t in $f_{t-1}(\theta_0)$. Adopting this assumption in the present case as well we need to truncate $g_{t-1}(\theta_0, \delta_0, \kappa_0)$, for example by

$$\begin{aligned} \overline{g}_0(\theta_0, \delta_0, \kappa_0) &= f_0(\theta_0), \\ \overline{g}_{t-1}(\theta_0, \delta_0, \kappa_0) &= f_{t-1}(\theta_0) + (\kappa_0 - \delta_0)\sum_{j=0}^{t-2}\delta_0^j(Y_{t-1-j} - f_{t-2-j}(\theta_0)) \\ \text{for } t &\geq 2. \end{aligned}$$

Since $g_{t-1}(\theta, \delta, \kappa) - \overline{g}_{t-1}(\theta, \delta, \kappa) \to 0$ exponentially for $t \to \infty$, using $\overline{g}_{t-1}(\theta, \delta, \kappa)$ as a proxy for $g_{t-1}(\theta, \delta, \kappa)$ yields, loosely speaking, asymptotically the same results as if the $g_{t-1}(\theta, \delta, \kappa)$'s were used.

More generally, suppose that $f_{t-1}(\theta)$ depends on the entire sequence $\{Z_{t-j}\}_{j=1}^{\infty}$, and that we can specify a truncated version $f_{t-1}^{\top}(\theta)$ of $f_{t-1}(\theta)$ that only depend on the observed lagged Y_t and X_t. One can formulate conditions regarding the speed of convergence to zero for $t \to \infty$ of $\sup_{\theta\in\Theta}|f_{t-1}(\theta) - f_{t-1}^{\top}(\theta)|$, $\sup_{\theta\in\Theta}||(\partial/\partial\theta')f_{t-1}(\theta) - (\partial/\partial\theta')f_{t-1}^{\top}(\theta)|$ and $\sup_{\theta\in\Theta}||(\partial/\partial\theta)(\partial/\partial\theta')f_{t-1}(\theta) - (\partial/\partial\theta)(\partial/\partial\theta')f_{t-1}^{\top}(\theta)||$ such that, after some modifications of the assumptions to take into account that $f_{t-1}^{\top}(\theta)$ is no longer strictly stationary, the above results carry over, with $f_{t-1}(\theta)$ replaced by $f_{t-1}^{\top}(\theta)$. This problem is related to the initial value problem discussed before. The details are left to the reader as an exercise.

10. Implementation

Choosing

$$\Upsilon = \times_{i=1}^{s+1}[-c, c]$$

for some $c > 0$, the WICM test statistic can be written in slightly different but asymptotically equivalent form as

$$\begin{aligned}\widehat{T}_n^{(1)} &= \sum_{k=1}^{L_n}\gamma_k\int_{\times_{i=1}^{(s+1).k}[-c,c]}\left|\frac{1}{\sqrt{n-L_n}}\sum_{t=L_n+1}^{n}\widehat{U}_t\exp\left(\mathbf{i}.\sum_{j=1}^{k}\tau_j'\widetilde{Z}_{t-j}\right)\right|^2 \\ &\quad\times \mathrm{d}\mu(\tau_1)\mathrm{d}\mu(\tau_2)...\mathrm{d}\mu(\tau_k)\end{aligned}$$

where as before, the $\widehat{U}_t$'s are the NLLS residuals, the $\widetilde{Z}_t$'s are vectors in $\mathbb{R}^{s+1}$ of bounded transformations of the standardized dependent and exogenous variables, and L_n is a subsequence of the effective sample size n. Throughout it will be assumed that μ is the uniform probability measure on Υ. The time shift accommodates the standardization procedure as well as the initial value issue.

Recall that under H_0, $\widehat{T}_n^{(1)} \xrightarrow{\mathrm{d}} T$, and that the actual WICM test is

$$\widehat{T}_n = \widehat{T}_n^{(1)}/\widehat{T}_n^{(2)},$$

where under H_0, $\widehat{T}_n^{(2)}$ is a consistent estimator of $E[T]$, and under H_1, $p\lim_{n\to\infty}\widehat{T}_n^{(2)} \in (0,\infty)$.

Lemma 10.1. *The statistic* $\widehat{T}_n^{(1)}$ *has the closed form expression*

$$\widehat{T}_n^{(1)} = \left(\sum_{k=1}^{L_n}\gamma_k\right)\left(\frac{1}{n-L_n}\sum_{t=L_n+1}^{n}\widehat{U}_t^2\right)$$

$$+2\frac{1}{n-L_n}\sum_{t_1=L_n+2}^{n}\sum_{t_2=L_n+1}^{t_1-1}\widehat{U}_{t_1}\widehat{U}_{t_2}P_{t_1,t_2}$$

where

$$P_{t_1,t_2}=\sum_{k=1}^{L_n}\gamma_k\prod_{j=1}^{k}\prod_{i=1}^{s+1}\frac{\sin\left(c\left(\widetilde{Z}_{i,t_1-j}-\widetilde{Z}_{i,t_2-j}\right)\right)}{c\left(\widetilde{Z}_{i,t_1-j}-\widetilde{Z}_{i,t_2-j}\right)}$$

with $\widetilde{Z}_{i,t}$ *component* i *of* $\widetilde{Z}_t$. *Moreover, the estimator* $\widehat{T}_n^{(2)}$ *has the closed form expression*

$$\widehat{T}_n^{(2)}=\left(\sum_{k=1}^{L_n}\gamma_k\right)\widehat{\sigma}_n^2-2.\mathrm{trace}\left[\widehat{A}_{1,n}^{-1}\widehat{C}_{2,n}\right]+\mathrm{trace}\left[A_{1,n}^{-1}\widehat{A}_{2,n}A_{1,n}^{-1}\widehat{C}_{1,n}\right],$$

where

$$\begin{aligned}
\widehat{\sigma}_n^2 &= \frac{1}{n-L_n}\sum_{t=L_n+1}^{n}\widehat{U}_t^2,\\
\widehat{A}_{1,n} &= \frac{1}{n-L_n}\sum_{t=L_n+1}^{n}\nabla f_t(\widehat{\theta}_n)\nabla f_t(\widehat{\theta}_n)',\\
\widehat{A}_{2,n} &= \frac{1}{n-L_n}\sum_{t=L_n+1}^{n}\widehat{U}_t^2\nabla f_t(\widehat{\theta}_n)\nabla f_t(\widehat{\theta}_n)',\\
\widehat{C}_{1,n} &= \frac{1}{(n-L_n)^2}\sum_{t_1=L_n+1}^{n}\sum_{t_2=L_n+1}^{n}\nabla f_{t_1}(\widehat{\theta}_n)\nabla f_{t_2}(\widehat{\theta}_n)'P_{t_1,t_2},\\
\widehat{C}_{2,n} &= \frac{1}{(n-L_n)^2}\sum_{t_1=L_n+1}^{n}\sum_{t_2=L_n+1}^{n}\widehat{U}_{t_1}^2\nabla f_{t_1}(\widehat{\theta}_n)\nabla f_{t_2}(\widehat{\theta}_n)'P_{t_1,t_2}.
\end{aligned}$$

Proof. Similar to Lemma 5.2 in the addendum to Bierens (1982) in Chapter 2. ■

Finally, it is not too hard, but rather tedious, to show that

Lemma 10.2. *A typical bootstrap version* $\widetilde{T}_n^{(1)}$ *of* $\widehat{T}_n^{(1)}$ *takes the form*

$$\widetilde{T}_n^{(1)}=\left(\sum_{k=1}^{L_n}\gamma_k\right)\frac{1}{n-L_n}\sum_{t=L_n+1}^{n}\varepsilon_t^2\widehat{U}_t^2$$

$$
\begin{aligned}
&+2\frac{1}{n-L_n}\sum_{t_1=L_n+2}^{n}\sum_{t_2=L_n+1}^{t_1-1}\varepsilon_{t_1}\varepsilon_{t_2}\widehat{U}_{t_1}\widehat{U}_{t_2}P_{t_1,t_2} \\
&-2\widetilde{d}_n'\widehat{A}_{1,n}^{-1}\widetilde{e}_n+\widetilde{e}_n'\widehat{A}_{1,n}^{-1}\widehat{C}_{1,n}\widehat{A}_{1,n}^{-1}\widetilde{e}_n
\end{aligned}
$$

where the ε_t's are i.i.d. $N(0,1)$, $\widehat{A}_{1,n}$ and $\widehat{C}_{1,n}$ are the same as before, and

$$
\begin{aligned}
\widetilde{e}_n &= \frac{1}{\sqrt{n-L_n}}\sum_{t=L_n+1}^{n}\varepsilon_t\widehat{U}_t\nabla f_t(\widehat{\theta}_n), \\
\widetilde{d}_n &= \frac{1}{(n-L_n)\sqrt{n-L_n}}\sum_{t_1=L_n+1}^{n}\sum_{t_2=L_n+1}^{n}\varepsilon_{t_1}\widehat{U}_{t_1}\nabla f_{t_2}(\widehat{\theta}_n).P_{t_1,t_2}.
\end{aligned}
$$

Since $\widetilde{T}_n^{(1)} \xrightarrow{\mathrm{d}} T^*$, where under H_0, $T^* \sim T$, and in general $p\lim_{n\to\infty}\widehat{T}_n^{(2)} = E[T^*]$, we may use $\widetilde{T}_n = \widetilde{T}_n^{(1)}/\widehat{T}_n^{(2)}$ as a bootstrap version of $\widehat{T}_n = \widehat{T}_n^{(1)}/\widehat{T}_n^{(2)}$.

11. A numerical example

11.1. Fitting an AR(1) model to an MA(1) process

In order to verify the finite sample performance of the WICM test, consider the zero-mean Gaussian MA(1) process

$$
Y_t = U_t - \rho U_{t-1}, \text{ where } U_t \sim \text{i.i.d. } N(0,1),\ |\rho|<1, \tag{11.1}
$$

and suppose that it is incorrectly assumed that this is an AR(1) process i.e.,

$$
H_0 : E\left[Y_t|\mathcal{F}_{-\infty}^{t-1}\right] = \theta_1 Y_{t-1} + \theta_2 \text{ a.s.} \tag{11.2}
$$

for some $\theta = (\theta_1,\theta_2)'$, where $\mathcal{F}_{-\infty}^{t-1} = \sigma\left(\{Y_{t-j}\}_{j=1}^{\infty}\right) = \sigma\left(\{U_{t-j}\}_{j=1}^{\infty}\right)$. Note that

$$
\begin{aligned}
\theta_0 &= (\theta_{0,1},\theta_{0,2})' = \arg\min_{(\theta_1,\theta_2)'\in\mathbb{R}^2} E\left[(Y_t-\theta_1 Y_{t-1}-\theta_2)^2\right] \\
&= \arg\min_{(\theta_1,\theta_2)'\in\mathbb{R}^2} E\left[(U_t-(\rho+\theta_1)U_{t-1}+\rho\theta_1 U_{t-2}-\theta_2)^2\right] \\
&= \arg\min_{(\theta_1,\theta_2)'\in\mathbb{R}^2} \left(1+\rho^2+2\rho\theta_1+(1+\rho^2)\theta_1^2+\theta_2^2\right) \\
&= \left(-\rho/(1+\rho^2),0\right)',
\end{aligned}
$$

hence, the error term of the AR(1) model involved is

$$\begin{aligned} U_t^* &= Y_t + \frac{\rho}{1+\rho^2} Y_{t-1} \\ &= U_t - \frac{\rho^3}{1+\rho^2} U_{t-1} - \frac{\rho^2}{1+\rho^2} U_{t-2}. \end{aligned}$$

It is easy to verify that

$$E\left[U_t^*.Y_{t-m}\right] = \begin{cases} 0 & \text{for} \quad m \neq 2, \\ -\frac{\rho^2}{1+\rho^2} & \text{for} \quad m = 2, \end{cases}$$

so that U_t^* is independent of all but one lagged Y_t, with Y_{t-2} the exception. Moreover

$$E\left[U_t^*.U_{t-m}\right] = \begin{cases} -\rho^3/(1+\rho^2) & \text{for} \quad m = 1, \\ -\rho^2/(1+\rho^2) & \text{for} \quad m = 2, \\ 0 & \text{for} \quad m \geq 3, \end{cases}$$

hence by joint normality,

$$E\left[U_t^* | U_{t-1}, U_{t-2}, ..., U_{t-m}\right] = -\frac{\rho^3}{1+\rho^2} U_{t-1} - \frac{\rho^2}{1+\rho^2} U_{t-2}$$

for $m \geq 2$ and thus

$$\begin{aligned} E\left[U_t^* | \mathcal{F}_{-\infty}^{t-1}\right] &= \lim_{m\to\infty} E\left[U_t^* | U_{t-1}, U_{t-2}, ..., U_{t-m}\right] \\ &= -\frac{\rho^2}{1+\rho^2}\left(\rho U_{t-1} + U_{t-2}\right) \\ &= -\frac{\rho^3}{1+\rho^2} Y_{t-1} - \rho^2 \sum_{j=0}^{\infty} \rho^j Y_{t-2-j}. \end{aligned}$$

The latter equality follows from inverting the MA(1) process (11.1), yielding $U_t = \sum_{j=0}^{\infty} \rho^j Y_{t-j}$.

Furthermore, note that

$$E\left[\left(E\left[U_t^* | \mathcal{F}_{-\infty}^{t-1}\right]\right)^2\right] = \frac{\rho^4}{1+\rho^2} \tag{11.3}$$

The values of this expectation are listed in the following table, for $\rho = 0.2, 0.3, ..., 0.9$.

Table 11.1: Values of the expectation (11.3)

ρ	$\rho^4/(1+\rho^2)$	ρ	$\rho^4/(1+\rho^2)$
0.2	0.0015	0.6	0.0953
0.3	0.0074	0.7	0.1611
0.4	0.0221	0.8	0.2498
0.5	0.0500	0.9	0.3625

Clearly, for low values of ρ it will be difficult to detect misspecification of the AR(1) model in finite samples.

11.2. WICM test results

I have generated independently four time series Y_t, $t = 1, 2, ..., 500$, according to (11.1) for $\rho = 0.8$, $\rho = 0.6$, $\rho = 0.4$ and $\rho = 0.2$, respectively, and conducted for each of these four time series the WICM test for the (false) null hypothesis (11.2).

The WICM test has been implemented as in the previous section, with $c = 5$, $\gamma_k = (0.9)^k$ and $L_n = [\sqrt{n}]$, where $[x]$ denotes the largest integer $\leq x$. Thus, $L_n = 22$. Moreover, the bootstrap sample size is $M = 500$.

In the following table, WICM is the test statistic $\widehat{T}_n = \widehat{T}_n^{(1)}/\widehat{T}_n^{(2)}$ as defined in Lemma 10.1, $b(0.01)$, $b(0.05)$, $b(0.10)$ are the bootstrap critical values for significance levels 1%, 5% and 10%, respectively, based on the bootstrap versions $\widetilde{T}_n = \widetilde{T}_n^{(1)}/\widehat{T}_n^{(2)}$ of $\widehat{T}_n$, with $\widetilde{T}_n^{(1)}$ defined in Lemma 10.2, and BPV is the bootstrap p-value.

Table 11.2: WICM test results

ρ	WICM	$b(0.01)$	$b(0.05)$	$b(0.10)$	BPV
0.8	2.11	2.05	1.53	1.35	0.006
0.6	2.16	1.78	1.47	1.32	0.002
0.4	1.55	1.77	1.52	1.38	0.042
0.2	0.56	1.82	1.46	1.30	1.000

First note that the WICM test statistics are lower than each of the three the upper bounds in (5.9). Thus, the tentative conclusion is that the false null hypothesis is not rejected. However, the results demonstrate that the upper bounds (5.9) are way too conservative. They are only informative if they indicate rejection at the 1% or 5% significance levels, because then we may conclude without conducting the bootstrap that the null hypothesis will be rejected on the basis of the bootstrap results as well.

For $\rho = 0.8$ and $\rho = 0.6$ the false AR(1) null hypothesis is strongly rejected and even for $\rho = 0.4$ the null hypothesis is rejected at the 5% significance level. In view of (11.3) and Table 11.1 the result for $\rho = 0.2$ is not surprising. In this case we need a much longer time series to detect the misspecification..

It would have been much more informative if I had done a complete Monte Carlo analysis of the finite sample power of the WICM test for the above MA(1)-AR(1) set up. However, each bootstrap round took about two hours computing time, so that such a Monte Carlo analysis with only 100 replications, for example, will take about one month non-stop!

12. Concluding remarks

In this addendum I have brought my old paper B84 up-to-date, by

- deriving the actual limiting null distribution of the proposed WICM test,
- deriving much sharper upper bounds of the critical values than those in B84 based on Chebyshev's inequality for first moments,
- showing how to approximate the actual asymptotic critical values via a bootstrap method, and
- proposing a martingale difference preserving standardization of the lagged conditioning variables such that all the asymptotic properties of the WICM test carry over.

However, the WICM test requires to make a number of choices, namely regarding

1. the absolutely continuous (with respect to Lebesgue measure) probability measure μ on Υ,
2. the compact set Υ itself, and
3. the positive sequence $\{\gamma_k\}_{k=1}^{\infty}$.

Under the null hypothesis that the time series regression model is correctly specified these choices do not matter too much. However, they do affect the finite sample power of the test. But under the alternative hypothesis that the model is incorrectly specified we do not know how the correct model looks like, so that we cannot choose Υ, μ and $\{\gamma_k\}_{k=1}^{\infty}$ such that the finite sample power of the WICM is "optimal". However, there are some practical considerations as well as sub-optimal adaptive procedures for these choices, which I will now discuss.

12.1. The probability measure μ

Boning and Sowell (1999) have shown that with μ the uniform probability measure on Υ the ICM test in Bierens and Ploberger (1997) is optimal in the sense of having the greatest weighted average local power as defined in Andrews and Ploberger (1994). It seems likely that this result carries over to the WICM test, although I have not verified this conjecture. Also, with μ the uniform probability measure and Υ a hypercube the WICM test statistic has a closed form expression. Therefore, it is recommended to choose for μ the uniform probability measure on Υ.

12.2. The compact set Υ

In Bierens (1982) and B84 it was recommended to choose Υ around the origin of the Euclidean space involved. However, for linear and nonlinear regression models with a constant term (as is usual the case) it is well-know that the least squares residuals sum up to zero, regardless whether the model is correctly specified or not. Consequently, in this case the empirical process $\widehat{W}_{k,n}$ in (1.13) satisfies $\widehat{W}_{k,n}(0,0,....0) \equiv 0$. Therefore, it seems better to choose Υ away from the origin of its Euclidean space.

According to Theorem 1 in Bierens and Ploberger (1997) we may choose Υ anywhere in its Euclidean space, but the best place where depends obviously on the correct model under the alternative hypothesis. A possible solution is to make Υ dependent on a parameter vector ξ representing the location and scale of Υ. Denoting this space by $\Upsilon(\xi)$, the WICM statistic becomes

$$\widehat{T}_n(\xi) = \sum_{k=1}^{L_n} \gamma_k \int_{\Upsilon(\xi)^k} \left|\widehat{W}_{k,n}(\tau_1, \tau_2, ...\tau_k)\right|^2 \mathrm{d}\mu(\tau_1|\xi)\mathrm{d}\mu(\tau_2|\xi)...\mathrm{d}\mu(\tau_k|\xi),$$

where $\mu(\tau|\xi)$ is the uniform probability measure on $\Upsilon(\xi)$. As long as we confine ξ to a compact or finite set Ξ it can be shown, similar to the addendum to Bierens (1982) in Chapter 2, that under H_0,

$$\sup_{\xi\in\Xi} \widehat{T}_n(\xi) \xrightarrow{\mathrm{d}} \sup_{\xi\in\Xi} T(\xi),$$

where

$$T(\xi) = \sum_{k=1}^{\infty} \gamma_k \int_{\Upsilon(\xi)^k} |W_k(\tau_1, \tau_2, ...\tau_k)|^2 \, \mathrm{d}\mu(\tau_1|\xi)\mathrm{d}\mu(\tau_2|\xi)...\mathrm{d}\mu(\tau_k|\xi),$$

whereas under H_1,

$$\sup_{\xi \in \Xi} \widehat{T}_n(\xi)/n \xrightarrow{\mathrm{d}} \sup_{\xi \in \Xi} \sum_{k=1}^{\infty} \gamma_k \int_{\Upsilon(\xi)^k} |\psi_k(\tau_1, \tau_2, ...\tau_k)|^2 \, \mathrm{d}\mu(\tau_1|\xi)\mathrm{d}\mu(\tau_2|\xi)...\mathrm{d}\mu(\tau_k|\xi),$$

where $\psi_k(\tau_1, \tau_2, ...\tau_k)$ is defined by (1.10).

In principle we can approximate the critical values of $\sup_{\xi \in \Xi} T(\xi)$ by the bootstrap approach in Theorem 6.2, because this approach is based on bootstrap versions of $\widehat{W}_{k,n}(\tau_1, \tau_2, ...\tau_k)$ only.

In practice the exact computation of $\sup_{\xi \in \Xi} \widehat{T}_n(\xi)$ is too tedious a numerical exercise. However, if we choose for Ξ a finite set then the computation of $\max_{\xi \in \Xi} \widehat{T}_n(\xi)$ is feasible.

12.3. The sequence of weights

The ideal weight sequence $\{\gamma_k\}_{k=1}^{\infty}$ for the WICM test is such that under H_1, γ_k is maximal when η_k is maximal, where η_k is defined in (1.15). But we don't know the η_k's. However, what we can do is to make the γ_k's dependent on a parameter with sufficient wide range to control the k for which γ_k is maximal. For example, choose for γ_k the probability of the Poisson(ω) distribution for $k-1$, i.e.,

$$\gamma_k(\omega) = \exp(-\omega)\frac{\omega^{k-1}}{(k-1)!}, \ k \in \mathbb{N},$$

with ω confined to a compact or finite set Ω in $(0, \infty)$. Then the WICM test statistic takes the form

$$\widehat{T}_n(\omega) = \sum_{k=1}^{L_n} \gamma_k(\omega)\widehat{B}_{n,k}.$$

It is not too hard to verify that

$$\sup_{\omega \in \Omega} \widehat{T}_n(\omega) \xrightarrow{\mathrm{d}} \sup_{\omega \in \Omega} \sum_{k=1}^{\infty} \gamma_k(\omega) B_k \text{ under } H_0,$$

$$\sup_{\omega \in \Omega} \widehat{T}_n(\omega)/n \xrightarrow{\mathrm{p}} \sup_{\omega \in \Omega} \sum_{k=1}^{\infty} \gamma_k(\omega)\eta_k > 0 \text{ under } H_1.$$

However, it seems unlikely that in this case the upper bounds (5.9) of the critical values are applicable to $\sup_{\omega\in\Omega}\widehat{T}_n(\omega)$,[8] but the bootstrap procedure still works. Note that then for *each* bootstrap version $\widetilde{T}_n(\omega) = \sum_{k=1}^{L_n}\gamma_k(\omega)\widetilde{B}_{n,k}$ of $\widehat{T}_n(\omega)$ we need to use $\sup_{\omega\in\Omega}\widetilde{T}_n(\omega)$ as the bootstrap version of $\sup_{\omega\in\Omega}\widehat{T}_n(\omega)$, which may take a lot of computing time. Therefore this approach is only feasible in practice if Ω is chosen finite.

References

Andrews, D. W. K., and W. Ploberger (1994): "Optimal Tests When a Nuisance Parameter Is Present Only under the Alternative", *Econometrica* 62, 1383-1414.

Bierens, H. J. (1982): "Consistent Model Specification Tests", *Journal of Econometrics* 20, 105-134, reprinted in Chapter 2

Bierens, H. J. (1984): "Model Specification Testing of Time Series Regressions", *Journal of Econometrics* 26, 323-353, reprinted in Chapter 3.

Bierens, H. J. (2004): *Introduction to the Mathematical and Statistical Foundations of Econometrics*, Cambridge University Press.

Bierens, H. J., and W. Ploberger (1997): "Asymptotic Theory of Integrated Conditional Moment Tests", *Econometrica* 65, 1129-1151, reprinted in Chapter 5.

Bierens, H. J., and L. Wang (2012): "Integrated Conditional Moment Tests for Parametric Conditional Distributions", *Econometric Theory* 28, 328-362, reprinted in Chapter 6.

Billingsley, P. (1968): *Convergence of Probability Measures*, Wiley & Sons.

Boning, B. and F. Sowell (1999): "Optimality of the Integrated Conditional Moment Test", *Econometric Theory* 15, 710-719.

Chung, K. L. (1974): *A Course in Probability Theory*, Academic Press.

Jennrich, R. I. (1969): "Asymptotic Properties of Nonlinear Least Squares Estimators", *Annals of Mathematical Statistics* 40, 633-643.

McLeish, D. L. (1974): "Dependent Central Limit Theorems and Invariance Principles", *Annals of Probability* 2, 620-628.

[8] We still have

$$\frac{\sum_{k=1}^{\infty}\gamma_k(\omega)B_k}{\sum_{k=1}^{\infty}\gamma_k(\omega)E[B_k]} \leq \sup_{n\geq 1}\frac{1}{n}\sum_{i=1}^{n}\left(\varepsilon_i(\omega)\right)^2.$$

where for each ω the $\varepsilon_i(\omega)$'s are i.i.d. standard normal, but they may depend on ω.